Numerical methods for optimization in finance:

optimized hedges for options and

optimized options for hedging

Dissertation
zur Erlangung des Doktorgrades an der
Mathematisch-Naturwissenschaftlichen Fakultät
der Universität Augsburg/ an der
Universität Pierre und Marie Curie

vorgelegt von

Tobias Lipp

2012

Betreuer: Prof. Dr. Ronald H. W. Hoppe
Prof. Dr. Olivier Pironneau

Erstgutachter: Prof. Dr. Ronald H. W. Hoppe
Zweitgutachter: Prof. Dr. Michael Günther

Tag der mündlichen Prüfung: 10. Mai 2012

Acknowledgments

This thesis has been jointly supervised by Prof. Dr. Ronald H. W. Hoppe and Prof. Dr. Olivier Pironneau. I would like to express my sincere gratitude to them for their continuous support and interest in my work and in my progress. Their long experience in academia and their deep scientific insight have always been very helpful.

I am very thankful to my committee members Prof. Dr. Yves Achdou, Prof. Dr. Michael Günther and Prof. Dr. Ekkehard Sachs for their time to read my dissertation, and for their valuable comments and suggestions.

I am grateful to Dr. Grégoire Loeper for the cooperation on hedged Monte Carlo. I would also like to thank Prof. Dr. Yvon Maday, Dr. Harbir Antil, Dr. Grégoire Derveaux and Prof. Dr. Denis Talay.

Recognition goes to the administrative members at Laboratoire Jacques-Louis Lions (LJLL) of Pierre and Marie Curie University and to the administrative members at Lehrstuhl für Angewandte Analysis mit Schwerpunkt Numerische Mathematik (LAM) at Augsburg University.

I am grateful to the Konrad-Adenauer-Stiftung for providing me financial and ideational support. The Franco-German University provided me with further material support to cover the additional expenses of a jointly supervised PhD.

Special thanks go to my present and former colleagues and friends at LJLL and at LAM. Finally, I would like to thank my parents for their constant support.

Contents

General introduction

Financial derivatives bare risk. Market models have been developed with the objective to provide a foundation for their valuation, for determination of hedging strategies and for risk assessment. Market deregulation, technical progress, globalization and increasing economic efficiency have led to more complex financial market models and financial derivatives. For the valuation of financial derivatives and for the determination of (optimal) hedging strategies sophisticated numerical methods are required.

This dissertation contributes to *optimization in finance* through *numerical methods*. The input consists of two parts:

- *Optimized hedges for options* We propose a numerical method to compute a trading strategy for the hedging of a financial derivative with N hedging instruments. The underlying mathematical framework is local risk minimization in discrete time. The method combines Monte Carlo simulation with least squares regression in analogy to the method of Longstaff and Schwartz [LS98].

- *Optimized options for hedging* We propose an optimal control approach for the optimization of European double barrier basket options. The basket consists of two assets. The objective is to control the payoff and the rebate at the upper barrier such that the delta of the option is as close as possible to a predefined constant. This gives rise to a control constrained optimal control problem for the (two-dimensional) Black-Scholes equation with Dirichlet boundary control and finite time control.

Risk and hedging

Broadly speaking, risk is the potential that a chosen action leads to an undesired state. A unique definition is missing. In finance, risk arises typically if one makes an investment or holds a portfolio whose future value is not known. One may distinguish between downside risk if the value is less than expected and upside risk if the value is larger than expected. The hedging of a portfolio requires the identification of the risk sources.

Let Π_t be the value of a portfolio at time t where $0 \leq t \leq T$ and $T \in \mathbb{R}_+$. Furthermore, let F be a risk source with size F_t at time t, $0 \leq t \leq T$. The current value $\Pi_0 = \Pi_0(F_0)$ is known but not the future values $\Pi_t = \Pi_t(F_\tau; 0 \leq \tau \leq t)$, $0 < t \leq T$. Hedging Π refers in our understanding to any action which reduces the dependence/sensitivity of the portfolio value Π_t to F. Commonly, the action is the inclusion of financial products into the portfolio.

Sensitivity measures and neutral portfolios

Let Π be a portfolio which consists of a financial derivative with underlying asset S. The underlying is risky and has stochastic volatility σ. The size of S, σ at time t, $0 \leq t \leq T$, is S_t, σ_t. The portfolio value Π_t is stochastic and has functional representation $\Pi_t = \Pi_t(S_t, \sigma_t)$ where $\Pi_t : \mathbb{R}_+^2 \to \mathbb{R}$ is twice continuously differentiable in the first variable and one times continuously differentiable in the second variable, $0 \leq t \leq T$. Three of the most famous sensitivity measures are the Greeks

$$\Delta_t := \frac{d\Pi_t}{dS} \qquad \text{(delta)}$$
$$\Gamma_t := \frac{d^2\Pi_t}{dS^2} \qquad \text{(gamma)}$$
$$\kappa_t := \frac{d\Pi_t}{d\sigma} \qquad \text{(vega)}$$

where $0 \leq t \leq T$. More about sensitivity measures and Greeks can be found in [Wil07]. The portfolio is called

$$\begin{bmatrix} \text{delta} \\ \text{gamma} \\ \text{vega} \end{bmatrix} \text{ neutral at } (t, S, \sigma) \text{ if } \begin{bmatrix} \Delta_t(S,\sigma) = 0 \\ \Gamma_t(S,\sigma) = 0 \\ \kappa_t(S,\sigma) = 0 \end{bmatrix}.$$

The size of Δ_t, Γ_t and κ_t at (t, S, σ) is the delta, gamma and vega exposure of the portfolio at (t, S, σ), respectively. Reduction of delta, gamma and vega exposure refers to reduction in absolute value of delta, gamma and vega exposure, respectively. Figure 0.1 illustrates the sensitivity reduction by Δ-hedging. The Δ-hedged portfolio is Δ-neutral at time 0 and stock price S_0.

1. Optimal dynamic hedging

Let H be a hedging objective. That is a cash flow at time T which is unknown at time $t < T$. The cash flow is for instance the payoff of a European option with maturity T. There are M risk sources $F^1, \ldots, F^M$ with size $F_t^1, \ldots, F_t^M$ at time t, $0 \leq t \leq T$. An investor who has to provide H (at time T) and has lower risk aversion is tempted to hedge this position. Let V be a hedge portfolio. There are N hedging instruments $X^1, \ldots, X^N$ with size (price) X_t^n at time t where $0 \leq t \leq T$ and $1 \leq n \leq N$. At time t, $0 \leq t \leq T$, the portfolio value is V_t and the portfolio consists of ϑ_t^n risky assets X^n where $1 \leq n \leq N$ and η_t riskless assets with value 1. Basically, the problem is how to choose $(\eta, \vartheta) := (\eta_t, \vartheta_t)_{0 \leq t \leq T}$.

This problem has been extensively studied in the last decades and is still active research. The approaches which have been suggested differ in the conditions on

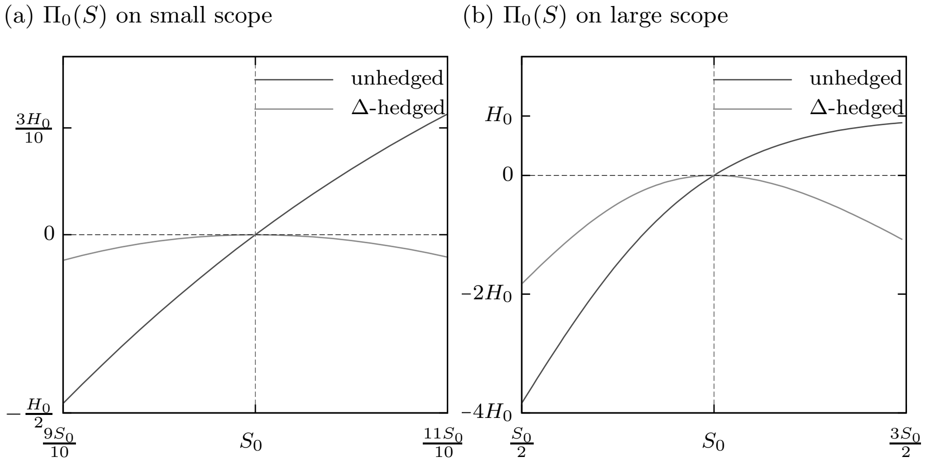


Figure 0.1.: In the unhedged case: $\Pi_0(S) := -1 \times H_0(S) + H_0(S_0) \times 1$, i.e. the portfolio consists of -1 options with value $H_0(S)$ and $H_0(S_0)$ bonds with value 1; in the Δ-hedged case: $\Pi_0(S) := -1 \times H_0(S) + \Delta(S_0)S + (H_0(S_0) - \Delta(S_0)S_0) \times 1$, i.e. the portfolio consists of -1 options with value $H_0(S)$, $\Delta(S_0)$ stocks with value S and $H_0(S_0) - \Delta(S_0)S_0$ bonds with value 1. The results have been obtained for the Black-Scholes model with parameters $r = 0.05$ (interest rate), $\sigma = 0.3$ (volatility), $T = 1$ (maturity), $K = 100$ (strike), $S_0 = 100$ and payoff function $\varphi(x) = \max(K - x, 0)$.

the parameters, the set of admissible strategies and the criteria which determine (optimal) (η, ϑ).

One of the first, simplest and most prominent approach is the Black-Scholes methodology [BS73] which is actually a pricing concept. The hedging strategy is a by-product. In the classical model the hedging objective is a European option with underlying risky asset. The risky asset is the only hedging instrument and the only risk source. The hedge is perfect in the sense that no residual risk is left over. The Black-Scholes model is what has been later called complete. In a complete model, any contingent claim is attainable/redundant by definition. This means that any contingent claim can be replicated by a self-financing hedging strategy [HK79, HP81]. A strategy (η, ϑ) is called replicating if the associated portfolio value is such that $V_T = H$. A strategy (η, ϑ) is called self-financing if $V_t = V_0 + \int\limits_0^t \vartheta_\tau dX_\tau$ for $0 \leq t \leq T$.

Non redundant contingent claims exist in incomplete market models. In this case a hedging strategy which provides no arbitrage opportunity is either self-financing or replicating but not both. A class of hedging concepts for incomplete market models is that of quadratic hedging. The two main representatives are mean-variance hedging and (local) risk minimization. Optimal strategies are self-financing in the mean-variance framework and replicating in the risk minimization framework. The risk minimization approach goes back to the diploma thesis of Schweizer [Sch84] and the paper of Föllmer and Sondermann [FS86]. Optimal strategies are determined by minimization of a quadratic risk measure. This measure is defined as the squared cost increment of the hedging strategy. In the sequel the risk minimization approach has been extended to the semi-martingale case [Sch88, Sch90, Sch91] (local risk minimization) and to discrete time hedging [SH88].

In the Black-Scholes model, the hedging strategy is explicitly given by a formula which can be evaluated easily. In more general frameworks, the existence of an (optimal) hedging strategy is sometimes hard to verify and there is often no explicit formula for it. For practical purpose, however, it is important being capable to compute the strategy. This can be the evaluation of a formula or the approximation via a numerical method. For local risk minimization in discrete time an explicit formula exists if there is one hedging instrument ($N = 1$). For general N, to our knowledge, there is no explicit formula and there is no numerical method available which can be used to approximate the hedging strategy. We propose a method for the general case in this dissertation. The method is called hedged Monte Carlo.

The basic principle of the hedged Monte Carlo method

Let $0 := t_0 < \ldots < t_k < \ldots < t_K = T$, $K \in \mathbb{N}$, be a finite number of time instances. The hedge portfolio can only be rebalanced at these dates. The hedging strategy is hence dynamic in discrete time. The hedging objective is an option with underlying asset S and payoff H. The available hedging instruments are the underlying S and financial derivatives Y - we set $X = (S, Y)$. At time t_k the underlying has price S_k and volatility σ_k - we set $F = (S, \sigma)$. The price of the derivative Y is Y_k at time t_k. It is assumed that Y_k depends on S_k and σ_k, $Y_k = Y_k(S_k, \sigma_k)$. The hedge portfolio consists of η units of the riskless asset, ϑ^1 units of the underlying asset (S) and ϑ^2 derivatives Y. At time t_k, $k = 0, \ldots, K$, the portfolio value is

$$V_k := \eta_k + \vartheta^1_k S_k + \vartheta^2_k Y_k.$$

The local risk minimization approach amounts to consider only hedging strategies which satisfy $V_K = H$ and $(\vartheta^1_0, \vartheta^2_0) = (0, 0)$. A locally risk-minimizing strategy $(V, \vartheta^1, \vartheta^2)$, $\eta_k = V_k - \vartheta^1_k S_k - \vartheta^2_k Y_k$, is then determined recursively backward in time for $k = K - 1, \ldots, 0$: find $(V_k, \vartheta^1_{k+1}, \vartheta^2_{k+1})$, such that

$$< \left[V_{k+1} - V_k - \vartheta^1_{k+1} \Delta S_{k+1} - \vartheta^2_{k+1} \Delta Y_{k+1}\right]^2 > \tag{1.1}$$

is minimized where $< \cdot >$ denotes the statistical average and $\Delta S_{k+1} := S_{k+1} - S_k$, $\Delta Y_{k+1} := Y_{k+1} - Y_k$ are the price increments of the hedging instruments.

In general it is not possible to solve problem (1.1) explicitly. In order to determine $(V_k, \vartheta^1_{k+1}, \vartheta^2_{k+1})$ numerically we proceed similarly as Longstaff and Schwartz [LS98].

They combined Monte Carlo simulation and least squares regression for the valuation of American options. We simulate (S_k, σ_k) and regress V_k, ϑ^1_{k+1} and ϑ^2_{k+1} over a set of basis functions

$$\{b_{k,d} = b_{k,d}(s,\sigma) \,|\, d = 1, \ldots, D_k\}, \quad k = K-1, \ldots 0$$

where $D_k \in \mathbb{N}$. That means V_k, ϑ^1_{k+1} and ϑ^2_{k+1} are approximated by

$$\begin{aligned}
V^D_k(S_k, \sigma_k) &:= \sum_{d=1}^{D_k} a^V_{k,d}\; b_{k,d}(S_k, \sigma_k) \\
\vartheta^D_{1,k+1}(S_k, \sigma_k) &:= \sum_{d=1}^{D_k} a^\vartheta_{1,k+1,d}\; b_{k,d}(S_k, \sigma_k) \\
\vartheta^D_{2,k+1}(S_k, \sigma_k) &:= \sum_{d=1}^{D_k} a^\vartheta_{2,k+1,d}\; b_{k,d}(S_k, \sigma_k)
\end{aligned}$$

with coefficients $a^V_{k,d}$, $a^\vartheta_{1,k+1,d}$, $a^\vartheta_{2,k+1,d} \in \mathbb{R}$, $d = 1, \ldots, D_k$.

Let $(S^{(i)}_k, \sigma^{(i)}_k)$, $i = 1, \ldots, I$, denote $I \in \mathbb{N}$ realizations of (S_k, σ_k), $k = K-1, \ldots, 0$. Then the coefficients are determined such that

$$\begin{aligned}
\sum_{i=1}^{I} \Bigg[V^D_{k+1}(S^{(i)}_{k+1}, \sigma^{(i)}_{k+1}) &- \sum_{d=1}^{D_k} a^V_{k,d}\, b_{k,d}(S^{(i)}_k, \sigma^{(i)}_k) \\
&- \sum_{d=1}^{D_k} a^\vartheta_{1,k+1,d}\, b_{k,d}(S^{(i)}_k, \sigma^{(i)}_k)\, \Delta S^{(i)}_{k+1} \\
&- \sum_{d=1}^{D_k} a^\vartheta_{2,k+1,d}\, b_{k,d}(S^{(i)}_k, \sigma^{(i)}_k)\, \Delta Y^{(i)}_{k+1} \Bigg]^2
\end{aligned} \tag{1.2}$$

is minimized where $\Delta S^{(i)}_{k+1} := S^{(i)}_{k+1} - S^{(i)}_k$ and $\Delta Y^{(i)}_{k+1} := Y^{(i)}_{k+1} - Y^{(i)}_k$. An approximation to the optimal hedging strategy is then obtained by setting $V^D_K = H$ and solving (1.2) recursively backward in time.

2. Optimal control of options

Exotic options and plain vanilla options are often considered as complementary. Plain vanilla options have a simple payoff function - in general the payoff depends only on the value of the underlying asset at exercise time. They are commonly exchange-traded. Exotic options are in contrast mostly found in the over-the-counter market. They have a more complex payoff function [CS97, JYC03, Nel99, Zha97] and are frequently tailored to the clients' specific needs. Two popular types of exotic options are basket options and barrier options. Basket options are options with two or more underlying assets[1]. They are often used to hedge portfolio risk. It is in general much less expensive to buy a basket option than to buy options on each single component of the portfolio. Barrier options are options which can be either activated (knock in options) or deactivated (knock out options) when the value of the underlying reaches some predetermined upper and/or lower bound (barrier). When deactivated they become either worthless or some fixed rebate is paid. For this reason they are in general less expensive than their counterparts without barrier.

An (institutional) investor holds a portfolio Π which consists of $\alpha = (\alpha_1, \alpha_2)$ stocks $S = (S_1, S_2)$. He expects that the stock prices will decrease in the near future but he does not want to or is not allowed to sell his stocks. The positions $\alpha = (\alpha_1, \alpha_2)$ might be that large that he will push the stock prices down if he sells the stocks, there could be tax issues or he is obligated to hold the stocks for a certain period etc. The value of the portfolio at time t is

$$\Pi_t = \alpha_1 S_{1,t} + \alpha_2 S_{2,t}$$

where $0 \leq t \leq T$ and $T \in \mathbb{R}_+$. The delta of the portfolio is

$$\nabla \Pi_t = \Delta := \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}, \quad t \in [0, T].$$

The delta is constant. To offset the delta one has to include a position with delta $-\Delta$ into the portfolio. One way to achieve this is to sell futures on S_1 and on S_2 but this has several drawbacks: i) there is a requirement to provide an initial margin and a safety margin to the clearing agency ii) additional margins have to be paid if the implied stock volatilities increase iii) the protection is against decreasing and increasing stock prices and hence one does not participate in the case the stock price increases. Selling call options might be an alternative. Standard call options, however, have no constant delta and this would require to adjust the position continuously. It would be more appealing to sell call options which have (almost) constant delta. But, how does such an option has to look like? We formulate the problem as a control constrained optimal control problem for the (two-dimensional) Black-Scholes equation with Dirichlet boundary control and finite time control (cf. below and part II).

The optimal control approach

The option prototype is a European double barrier knock-out call on the basket $\{S_1, S_2\}$. The payoff function and the rebate at the upper barrier are parametrized. These parameters are then the control variable. At the lower barrier the rebate is zero.

Optimal control problems are about finding a control variable u that minimizes a cost functional J. Given a control u there is an associated state variable $y = y(u)$

[1]The underlying can equally be an index or an exchange rate.

which typically has to satisfy one or several differential or algebraic equations. Besides, there can be further (inequality) constraints on the control and/or on the state. The cost functional J is in general a function of both the state and the control variable i.e. $J = J(y, u)$. It is sometimes written in control reduced form ($J = J(u)$) or in state reduced form ($J = J(y)$). Optimal control problems are classified according to the form of their cost functionals, according to their equality and inequality constraints and according to the domain of the controls variables. An introductory text to optimal control is [Trö10, HPUU08]. Further information can be found in [Lio71].

In our case the control is a vector of parameters $u = (u_1, \ldots, u_M)^T \in \mathbb{R}^M$ where $M \in \mathbb{N}$. The control governs the payoff function $g = g(u)$. The rebate at the upper barrier is u_M. There are control constraints

$$u_{min,i} \leq u_i \leq u_{max,i} \quad , \quad i = 1, \ldots, M \tag{2.1}$$

with $u_{min} = (u_{min,1}, \ldots, u_{min,M})^T \in \mathbb{R}^M$ and $u_{max} = (u_{max,1}, \ldots, u_{max,M})^T \in \mathbb{R}^M$. The cost functional is defined by

$$J(y_Q, u) := \frac{1}{2} \int_0^T \int_\Omega |\nabla y_Q - d|^2 dS dt,$$

where the state y_Q , $Q := \Omega \times (0, T)$, denotes the price of the option and $d = (d_1, d_2)^T \in \mathbb{R}^2$ is the desired gradient (i.e. in financial terms the desired delta). The option reaches maturity at time T where $T \in \mathbb{R}_+$. Hence the first integral is over the life time of the option. The second integral is over the domain between the barriers

$$\Omega := \{S = (S_1, S_2) \in \mathbb{R}^2_+ \mid K_{min} < |S| < K_{max}\}$$

where K_{min}, $K_{max} \in \mathbb{R}_+$ such that $K_{min} < K_{max}$ and $|S| := S_1 + S_2$. The boundaries are

$$\begin{aligned} \Gamma_1 &:= (K_{min}, K_{max}) \times \{0\}, & \Gamma_2 &:= \{0\} \times (K_{min}, K_{max}), \\ \Gamma_3 &:= \{S \in \mathbb{R}^2_+ \mid |S| = K_{min}\}, & \Gamma_4 &:= \{S \in \mathbb{R}^2_+ \mid |S| = K_{max}\}. \end{aligned}$$

The lower barrier of the option is Γ_3 and the upper barrier is Γ_4, see figure 2.1. The state equation is a boundary value problem for the two-dimensional Black-Scholes pde

$$\frac{\partial y_Q}{\partial t} + L_\Omega(t) y_Q = 0 \qquad \text{in } Q, \tag{2.2a}$$

$$y_Q = y_{\Sigma_j} \qquad \text{on } \Sigma_j := \Gamma_j \times (0, T) \ , \ 1 \leq j \leq 4, \tag{2.2b}$$

with final time condition

$$y_Q(\cdot, T) = y_{Q,T} \qquad \text{in } \Omega \tag{2.2c}$$

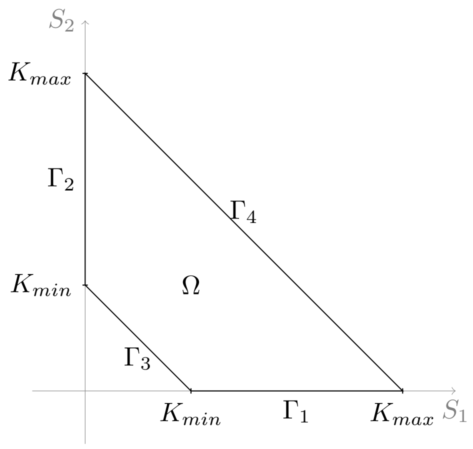


Figure 2.1.: The domain Ω with knock-out barriers Γ_3 and Γ_4.

where $L_\Omega(t)$, $t \in [0, T]$, stands for the Black-Scholes operator, a second order elliptic operator, y_{Σ_j}, $1 \le j \le 4$, are boundary functions and $y_{Q,T}$ is the final time function. We set

$$y_{Q,T} = g(u), \qquad y_{\Sigma_3} = 0, \qquad y_{\Sigma_4} = u_M.$$

The boundary functions y_{Σ_j}, $j = 1, 2$, are the solution of boundary value problems for one-dimensional Black-Scholes pde analog to (2.2).

Particular features of the optimal control problem are the one-dimensional problems for the Dirichlet boundary functions y_{Σ_j}, $j = 1, 2$, the geometry of the computational domain Ω and the final time condition (2.2c). Furthermore the cost functional J is defined without regularization term and the optimization object is the gradient of the state and not the state itself.

We will replace time by time to maturity and consider the weak formulation of the optimal control problem in a weighted Sobolev space setting. This gives rise to a non-symmetric bilinear form. It will be shown that the optimal control problem admits a unique solution. The rest of our analysis concerns the solution of the optimal control problem. We derive necessary optimality conditions based on the existence of an adjoint state p satisfying a parabolic final time problem and due to the constraints on the control (2.1) a variational inequality.

Standard methods for the spatial discretization of pdes are the finite difference method, the finite element method and the finite volume method. We will use the finite element method. A priori and a posteriori error estimates are available, cf. e.g. [AP05] and the references therein. For the discretization of the temporal domain we use the implicit Euler method. The method has consistency and convergence order 1. To our knowledge however, the convergence of the solution of the fully discrete optimal control problem to the solution of the original optimal control problem has not been established yet.

We will derive first order necessary optimality conditions for the fully discrete optimal control problem and present an algorithmic approach for their solution. The approach is based on the projected BFGS method gradient method with Armjio line search.

3. Outline of the thesis

The thesis is organized as follows.

Part I is on optimal dynamic hedging and consists of three chapters. In chapter 1 we formulate the hedge problem in the framework of local risk minimization and present for the discretization of the problem a numerical method (hedged Monte Carlo). The following two chapters concern the application of the hedged Monte Carlo method. In chapter 2 the problem is to hedge with vanilla put options and in chapter 3 the problem is to hedge with variance swaps. In chapter 2 the model features only one source of uncertainty: the price of the underlying of the hedging objective whereas in chapter 3 there are two sources of uncertainty: the price of the underlying of the hedging objective and its variance/volatility.

Part II is on optimal control of European double barrier basket options. It consists of one chapter.

Appendix B is about pricing European options on a basket of assets. We show how stochastic (Monte Carlo) and deterministic methods (quadrature and PDE methods) can be combined for this purpose. The ultimate objective has been to define mixed methods which provide speed-up with respect to standard methods. This work does not fit to the framework of *hedging and optimization*. For this reason it has been shifted to the appendix.

Part I.

Optimal dynamic hedging

The work on optimal dynamic hedging been supervised by Grégoire Loeper. I have received general advice from Olivier Pironneau.

1. The hedged Monte Carlo method

Abstract: The theory called local risk minimization provides a concept for the determination of optimal strategies for the hedging of financial derivatives with N hedging instruments. These strategies are in general not explicitly given and require numerical approximation. We present a discretization method for this purpose. The method gives rise to an efficient algorithm which allows to compute approximations to optimal strategies. We analyze the resultant error.

Contents

1.1. Introduction

In complete market models the price of a contingent claim is unique and there is a hedging strategy for the contingent claim which allows to eliminate the risk completely. Real markets however can be considered as incomplete: discontinuous price processes, existence of bid/ask spreads, limited liquidity, transaction costs, discrete hedging etc. In general, risk cannot be eliminated and the trader chooses the hedging strategy according to his risk aversion.
Transaction costs limit the rebalancing of the hedge portfolio to discrete points in time. A popular strategy is to adjust the hedge portfolio according to current Black-Scholes greeks. They can be computed easily using closed form expressions. Black-Scholes delta hedging in discrete time has been investigated in [BS72, BE80]. The strategy however has at least two drawbacks: the Black-Scholes model is often too far from reality and the strategy is only optimal in the Black-Scholes model where continuous hedging is possible.
If more sophisticated models and/or hedging approaches are considered, in particular if there is more than one hedging instrument, the (optimal) strategy is in general not given by a closed-form expression which can be evaluated with little effort. Numerical methods are required to approximate the optimal trading strategy in these cases. For practical importance the method should be easy to employ and it should give rise to an algorithm which is fast, i.e. for which the trade-off between precision and computing time is in good ratio.

Local risk minimization is such a hedging approach. The theory is quite well understood, [FS86, Sch88, SH88] to name a few contributions, but numerical results demonstrating its practicability are sparse. There are two papers of Heath, Platen and Schweizer [HPS01a, HPS01b] where local risk minimization is compared with mean-variance hedging[1]. They provided numerical results for the Stein-Stein [SS91] and the Heston model [Hes93] but only for the case if there is one hedging instrument, the underlying asset of the hedging objective. The numerical results have been obtained by a procedure which involves the following main steps: a) derivation of the minimal martingale measure; b) determination of the dynamics of the state variables under the minimal martingale measure; c) solution of a parabolic partial differential equation for the portfolio value on a domain with dimension equal to the number of state variables (if the model is Markov) Step a) and c) limit the applicability of the approach. Besides, the approach seems to be only admissible if there is exactly one hedging instrument. This is a severe drawback. Many trading desks hedge their risk with the underlying asset(s) and some other financial products.

In 2001 Potters, Bouchaud and Sestovic [PBS01] came up with a Monte Carlo method for pricing exotic options and for pricing options with early exercise feature. The method exhibits very low variance, since it makes implicitly use of the effect of introducing a control variate in Monte Carlo simulation. If X_k and X_{k+1} are the prices of the underlying at time k respectively at time $k+1$, the control variate

[1] Also known as total risk minimization.

is $\Delta X_{k+1} := X_{k+1} - X_k$. Given the price of the option at time $k+1$, V_{k+1}, the problem is to find (V_k, ϑ_{k+1}) such that

$$E\left[V_{k+1} - V_k - \vartheta_{k+1}\Delta X_{k+1}\right]^2 \tag{1.1.1}$$

is minimized where V_k is (by definition) the price of the option at time k. The optimal ϑ_{k+1} is a by-product of the method and it is actually the optimal hedge in the sense of local risk minimization.

In contrast to the approach of Heath et al., the minimal martingale measure does not have to be determined explicitly. The expectation in (1.1.1) is with respect to the real world probability measure. Instead of solving a partial differential equation for the portfolio value one has to solve least squares problems. The method shares the flavor of Longstaff and Schwartz's Monte Carlo method for the valuation of American options [LS98]. It is based on Monte Carlo simulation, the dynamic programming principle and least squares regression.

In §1.3 we introduce a discretization method similar to [PBS01, LS98] which allows to approximate the hedging strategy given N hedging instruments in the framework of local risk minimization. But before we formulate the problem in §1.2. The method gives rise to an algorithm described in §1.4. The error is then analyzed in §1.5.

1.2. Local risk minimization

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $\mathbb{F} = (\mathcal{F}_k)_{k=0}^K$, $K \in \mathbb{N}$, be a filtration with $\mathcal{F}_K = \mathcal{F}$ and $P(A) \in \{0,1\}$, $A \in \mathcal{F}_0$. Let $H : \Omega \to \mathbb{R}$ be a $\mathcal{F}$-measurable random variable and $X = (X_k)_{k=0}^K$ be a $\mathbb{F}$-adapted[2] $\mathbb{R}^N$-valued, $N \in \mathbb{N}$, stochastic process. The interpretation is that H describes the payoff of a financial derivative and that there are N (different) hedging instruments with price X_k at time k , $k = 0, \ldots, K$. All prices are quoted in terms of a numéraire asset. Next, we define what we understand under a hedging strategy.

Definition 1.2.1. i) A hedging strategy is a pair (η, ϑ) of two stochastic processes such that $\eta = (\eta_k)_{k=0}^K$ is $\mathbb{F}$-adapted and $\mathbb{R}$-valued and $\vartheta = (\vartheta_k)_{k=0}^K$ is $\mathbb{F}$-predictable[3] and $\mathbb{R}^N$-valued with $\vartheta_0 = 0$ P-a.s.
ii) The value process $V = (V_k)_{k=0}^K$ associated with a hedging strategy (η, ϑ) is defined by

$$V_k := \eta_k + \vartheta_k \cdot X_k, \qquad k = 0, \ldots, K. \tag{1.2.1}$$

Think of a portfolio built to hedge H. At time k the portfolio consists of η_k numéraire assets and $\vartheta_{n,k}$ hedging instruments of type X_n, $n = 1, \ldots, N$. We introduce another two processes which allow to monitor the (cumulated) gain/loss and the (cumulated) cost associated with a trading strategy.

[2] A stochastic process $Z = (Z_k)_{k=0}^K$ is called $\mathbb{F}$-adapted if Z_k is $\mathcal{F}_k$-measurable for $k = 0, \ldots, K$.
[3] A stochastic process $Z = (Z_k)_{k=0}^K$ is called $\mathbb{F}$-predictable if Z_k is $\mathcal{F}_{k-1}$-measurable for $k = 1, \ldots, K$ and Z_0 is $\mathcal{F}_0$-measurable.

Convention 1.2.2. For any stochastic process $Z = (Z_k)_{k=0}^K$ we denote by ΔZ_{k+1}, $k = 0, \dots, K-1$, the increment

$$\Delta Z_{k+1} := Z_{k+1} - Z_k.$$

Definition 1.2.3. Let (η, ϑ) be a hedging strategy.
i) The gain process $G = (G_k)_{k=0}^K$ of (η, ϑ) is defined by

$$G_0 := 0 \quad \text{and} \quad G_k := \sum_{\ell=1}^{k} \vartheta_\ell \cdot \Delta X_\ell, \quad k = 1, \dots, K. \tag{1.2.2}$$

ii) The cost process $C = (C_k)_{k=0}^K$ of (η, ϑ) is defined by

$$C_k := V_k - G_k, \qquad k = 0, \dots, K. \tag{1.2.3}$$

The central element of the local risk minimization approach is the definition of the risk process featuring the risk measure, the definition of the class of admissible hedging strategies and the definition/characterization of locally risk-minimizing strategies.
An admissible hedging strategy has to be H-replicating, i.e.

$$V_K = H \quad P\text{-a.s.} \tag{1.2.4}$$

but not necessarily self-financing, i.e. the cost process does not have to be constant. That means it is possible that

$$C_k \neq V_0 \quad P\text{-a.s.}$$

for some $k \in \{1, \dots, K\}$.

Optimality (risk minimality) is defined with respect to a quadratic measure of risk. The definition of the risk measure (process) requires square integrability of some of the quantities involved.

Assumption 1.2.4. It is assumed that H and X_k, $k = 0, \dots, K$ are square integrable with respect to P, i.e.

i) $H \in L^2(\Omega, \mathbb{R}, \mathcal{F}, P) =: L^2(P)$

ii) $X_k \in L^2(\Omega, \mathbb{R}^N, \mathcal{F}, P)$, $k = 0, \dots, K$.

Definition 1.2.5. A hedging strategy (η, ϑ) is called admissible (for H) if it satisfies (1.2.4) and if

a) $V_k \in L^2(P)$, $k = 0, \dots, K$

b) $\vartheta_k^n \Delta X_k^n \in L^2(P)$, $k = 1, \dots, K$, $n = 1, \dots, N$.

Definition 1.2.6. i) The risk process $R = (R_k)_{k=0}^{K-1}$ of an admissible strategy is defined by

$$R_k := E\left[\Delta C_{k+1}^2 | \mathcal{F}_k\right] \qquad k = 0, \dots, K-1. \tag{1.2.5}$$

ii) An admissible strategy $(\eta^{\mathtt{lrm}}, \vartheta^{\mathtt{lrm}})$ is called locally risk-minimizing if for $k = 0, \dots, K-1$ there holds

$$R_k(\eta^{\mathtt{lrm}}, \vartheta^{\mathtt{lrm}}) \leq R_k(\eta, \vartheta) \quad P\text{-a.s.}$$

for any admissible (η, ϑ) such that

$$V_{k+1} = V_{k+1}^{\mathtt{lrm}} \quad P\text{-a.s.}$$

Note, since

$$\begin{aligned} R_k &= E\left[(C_{k+1} - C_k)^2 | \mathcal{F}_k\right] \\ &= E\left[(V_{k+1} - V_k - \vartheta_{k+1} \cdot \Delta X_{k+1})^2 | \mathcal{F}_k\right], \qquad k = 0, \dots, K-1, \end{aligned}$$

the process R can be viewed as a measure of the (local) risk with R_k the up to time $k+1$ expected risk at time k.

The existence of a locally risk-minimizing strategy is in general not given. If $N = 1$, it can be shown that under an additional assumption[4] on X a locally risk-minimizing strategy exists [FS04, proposition 10.10]. For the general case, $N \geq 1$, the existence of locally risk-minimizing strategies can be characterized in terms of H: H has to admit some sort of orthogonal decomposition [FS04, corollary 10.14].

A characterization of locally risk-minimizing strategies is given in the following theorem.

Theorem 1.2.7. *An admissible strategy (η, ϑ) is locally risk-minimizing if and only if it is mean self-financing, i.e.*

$$E\left[\Delta C_{k+1} | \mathcal{F}_k\right] = 0 \qquad P\text{-a.s.} \tag{1.2.6}$$

for $k = 0, \dots, K-1$ and C is strongly orthogonal to X, i.e.

$$Cov(\Delta C_{k+1}, \Delta X_{k+1} | \mathcal{F}_k) = 0 \qquad P\text{-a.s.}$$

for $k = 0, \dots, K-1$.

The proof of the theorem, [FS04, proof of theorem 10.9], shows that a locally risk-minimizing strategy, if it exists, can be determined recursively backward in time. The recursion starts by setting $V_K^{\mathtt{lrm}} = H$. After that $(V_{K-1}^{\mathtt{lrm}}, \vartheta_K^{\mathtt{lrm}})$ is determined by minimizing

$$R_{K-1} = E\left[(V_K - V_{K-1} - \vartheta_K \cdot \Delta X_K)^2 | \mathcal{F}_{K-1}\right] \tag{1.2.7}$$

[4] If X has bounded mean-variance trade-off, i.e. if there is a constant c such that $(E[\Delta X_{k+1} | \mathcal{F}_k])^2 \leq c Var[\Delta X_{k+1} | \mathcal{F}_k]$ P-a.s., $k = 0, \dots, K-1$.

over all admissible strategies (η, ϑ) with $V_K = V_K^{\mathtt{lrm}}$ P-a.s. In view of definition (1.2.1), $\eta_K^{\mathtt{lrm}}$ is obtained by setting $\eta_K^{\mathtt{lrm}} = V_K^{\mathtt{lrm}} - \vartheta_K^{\mathtt{lrm}} \cdot X_K$. Next, $(V_{K-2}^{\mathtt{lrm}}, \vartheta_{K-1}^{\mathtt{lrm}})$ is determined by minimizing

$$R_{K-2} \;=\; E\left[(V_{K-1} - V_{K-2} - \vartheta_{K-1} \cdot \Delta X_{K-1})^2 | \mathcal{F}_{K-2}\right] \tag{1.2.8}$$

over all admissible strategies (η, ϑ) with $V_{K-1} = V_{K-1}^{\mathtt{lrm}}$ P-a.s. In turn, $\eta_{K-1}^{\mathtt{lrm}}$ is set to $V_{K-1}^{\mathtt{lrm}} - \vartheta_{K-1}^{\mathtt{lrm}} \cdot X_{K-1}$.
This shows that induction over time $k = K-1, \ldots, 0$ determines completely a locally risk-minimizing strategy $(\eta^{\mathtt{lrm}}, \vartheta^{\mathtt{lrm}})$. The strategy is in general not unique since there may be more than just one minimizer of

$$R_{k+1} = E\left[(V_{k+1} - V_k - \vartheta_{k+1} \cdot \Delta X_{k+1})^2 | \mathcal{F}_k\right]$$

under the admissible strategies (η, ϑ) with $V_{k+1} = V_{k+1}^{\mathtt{lrm}}$ P-a.s. However, all locally risk-minimizing strategies have the same risk process, cf. [FS04, corollary 10.14].

The recursive procedure adumbrated above is described in algorithm 1. Before we introduce some notation.

Definition 1.2.8. For $k = 0, \ldots, K-1$ let

$$\mathcal{A}_{0,k}^r := L^2(\Omega, \mathbb{R}, \mathcal{F}_k, P) \tag{1.2.9a}$$

$$\mathcal{A}_{n,k}^r := \{\, Z : \Omega \to \mathbb{R} \,|\, Z \ \mathcal{F}_k\text{-measurable},\ Z\Delta X_{k+1}^n \in L^2(P)\,\},\ n = 1, \ldots, N. \tag{1.2.9b}$$

Algorithm 1 Local risk minimization

1: set $V_K^{\mathtt{lrm}} = H$
2: **for** $k = K-1$ to 0 **do**
3: solve
4:
$$\min_{\substack{V_k \in \mathcal{A}_{0,k}^r,\\ \vartheta_{k+1}^n \in \mathcal{A}_{n,k}^r,\\ n=1,\ldots,N}} E\left[(V_{k+1}^{\mathtt{lrm}} - V_k - \vartheta_{k+1} \cdot \Delta X_{k+1})^2 | \mathcal{F}_k\right] \tag{1.2.10}$$
5: denote the solution by $(V_k^{\mathtt{lrm}}, \vartheta_{k+1}^{\mathtt{lrm}})$
6: **end for**
7: set $\vartheta_0^{\mathtt{lrm}} = 0$ and $\eta_0^{\mathtt{lrm}} = V_0^{\mathtt{lrm}} - \vartheta_0^{\mathtt{lrm}} \cdot X_0$

1.3. The hedged Monte Carlo method

Algorithm 1 describes formally how a locally risk-minimizing strategy can be constructed. The problem which arises first when the strategy has really to be constructed is not the computational complexity to solve the local risk minimization (lrm) problem (1.2.10). The problem is that the spaces $\mathcal{A}^r_{n,k}$, $n = 0, \dots, N$ are only formally known. The elements of $\mathcal{A}^r_{n,k}$, $n = 0, \dots, N$, are random variables defined on Ω. While $\mathcal{A}^r_{0,k}$ is an L^2-space the elements of $\mathcal{A}^r_{n,k}$, $n = 1, \dots, N$, are in general not square integrable. In this section a discrete counterpart to problem (1.2.10) is derived. The way of discretization is referred to as the hedged Monte Carlo method in later sections/chapters.

Step 1: Reformulation of the lrm problem

Assumption 1.3.1. It is assumed that

$$\mathcal{F}_k = \sigma\left(\{\, F_{m,\ell} \,|\, \ell = 0, \dots, k;\ m = 1, \dots, M \,\}\right),$$

where $F = (F_k)_{k=0}^K$ is a $\mathbb{R}^M$-valued stochastic process and $M \in \mathbb{N}$.

Assumption 1.3.1 has been introduced for two reasons: a) the filtration $\mathcal{F}$ got herewith a clear structure b) with F and X we can distinguish between 'information' and the price of the hedging instruments. To model the problem described in the general introduction (§1) we would set $M = 2$ and identify $F = (F^1, F^2)$ with (S, σ) and $X = (X^1, X^2)$ with (S, Y).

Since F generates $\mathbb{F} = (\mathcal{F}_k)_{k=0}^K$, any $\mathcal{F}_k$-measurable random variable can be represented as a function of $F_0, \dots, F_k$, in particular $X_k = X_k(F_0, \dots, F_k)$, cf. e.g. [McL05]. This allows us to reformulate problem (1.2.10).

Definition 1.3.2. For $k = 0, \dots, K-1$ let

$$\mathcal{A}^f_{n,k} \ :=\ \{\, Z : F_{0\dots k}(\Omega) \to \mathbb{R} \,|\, Z(F_{0\dots k}) \in \mathcal{A}^r_{n,k} \,\},\ \ n = 0, \dots, N, \tag{1.3.1}$$

where $F_{0\dots k} := (F_0, \dots, F_k)$ and $F_{0\dots k}(\Omega) := (F_0, \dots, F_k)\,(\Omega) \subset \mathbb{R}^{(k+1)M}$ is the set of the paths up to time k.

In terms of $\mathcal{A}^f_{n,k}$, $n = 0, \dots, N$, problem (1.2.10) becomes

$$\min_{\substack{V_k \in \mathcal{A}^f_k,\\ \vartheta^n_{k+1} \in \mathcal{A}^f_{n,k},\\ n=1,\dots,N}} E\Bigg[\Big(V^{\mathtt{lrm}}_{k+1}(F_{0\dots k+1}) - V_k(F_{0\dots k}) \\ -\vartheta_{k+1}(F_{0\dots k}) \cdot \Delta X_{k+1}(F_{0\dots k+1})\Big)^2 \Big| \mathcal{F}_k\Bigg] \tag{1.3.2}$$

where $\Delta X_{k+1}(F_{0\dots k+1}) := X_{k+1}(F_{0\dots k+1}) - X_k(F_{0\dots k})$, $0 \le k \le K-1$.

Since the value of F_0 is known with probability 1 at time 0, we may assume without loss of generality that $F_0(\Omega)$ consists of a single element.

Assumption 1.3.3. It is assumed that

$$\mathrm{card}\,(F_0(\Omega)) = 1.$$

Consequently $\mathcal{A}^f_{n,0}$, $n = 0, \dots, N$, are spanned by a constant and have dimension 1. For $k > 0$, the spaces $\mathcal{A}^f_{n,k}$, $n = 0, \dots, N$, are like $\mathcal{A}^r_{n,k}$, $n = 0, \dots, N$, only formally known. It is known that $F_{0\dots k}(\Omega) \subset \mathbb{R}^{(k+1)M}$ but in general not more. The elements of $\mathcal{A}^f_{n,k}$, $n = 0, \dots, N$, are not square integrable in general.

Step 2: Semi discrete formulation

We introduce some notation to replace $\mathcal{A}^f_{n,k}$, $n = 0, \dots, N$, in a subsequent step.

Definition 1.3.4. For $k = 0, \dots, K-1$ let $D_k \in \mathbb{N}$ and

$$A_k := \operatorname{span}\{b_{k,1}, \dots, b_{k,D_k}\},$$

where

$$b_{k,j} : U_k \to \mathbb{R}, \quad j = 1, \dots, D_k$$

are (basis) functions defined on sets $U_k \subset \mathbb{R}^{M_k}$, $M_k \leq (k+1)M$. Let

$$f_k : F_{0\dots k}(\Omega) \subset \mathbb{R}^{(k+1)M} \to U_k$$

be some additional functions.

A semi-discrete formulation of (1.3.2) is

$$\min_{\substack{V_k \in A_k, \\ \vartheta^n_{k+1} \in A_k, \\ n=1,\dots,N}} E\Bigg[\Big(V^{\mathtt{lrm}}_{k+1}(f_{k+1}(F_{0\dots k+1})) - V_k(f_k(F_{0\dots k})) - \vartheta_{k+1}(f_k(F_{0\dots k})) \cdot \Delta X_{k+1}(F_{0\dots k+1})\Big)^2 \Big| \mathcal{F}_k\Bigg], \tag{1.3.3}$$

where $0 \leq k \leq K-1$.

Remark 1.3.5. i) In this section the discretization of the lrm problem is described formally. The parameters D_k, $b_{k,j}$ and f_k are chosen later.
ii) Basis functions of $(k+1)M$ variables may impose a computational problem due to the size of $(k+1)M$. In a Markovian setting one may set $f_k(F_{0\dots k}) := F_k$ and in another setting one may (have to) omit some information. For this reason we will sometimes call the functions f_k *functions for model reduction* or alike.
iii) For notational simplicity we have not introduced spaces $A_{n,k}$ for $n = 0, \dots, N$.

Step 3: Fully discrete formulation

We discretize the conditional expectation ($E[\cdot|\mathcal{F}_k]$) in (1.3.3) by Monte Carlo simulation. Let $F^{(1)}, \dots, F^{(I)}$ be I realizations of F where $I \in \mathbb{N}$.
The fully discrete formulation of (1.3.2) is

$$\min_{\substack{V_k \in A_k, \\ \vartheta^n_{k+1} \in A_k, \\ n=1,\dots,N}} \frac{1}{I} \sum_{i=1}^{I} \Big(V^{\mathtt{lrm}}_{k+1}(f_{k+1}(F^{(i)}_{0\dots k+1})) - V_k(f_k(F^{(i)}_{0\dots k})) - \vartheta_{k+1}(f_k(F^{(i)}_{0\dots k})) \cdot \Delta X_{k+1}(F^{(i)}_{0\dots k+1})\Big)^2, \tag{1.3.4}$$

where $0 \leq k \leq K-1$.

1.4. The hedged Monte Carlo algorithm

A least squares problem is an optimization problem which arises when the solution of an overdetermined system is approximated by the method of least squares. Typically, a least squares problem has the form

$$\min_{x \in \mathbb{R}^\alpha} \|y - Bx\|_2^2 \,, \tag{1.4.1}$$

where $\|\cdot\|_2$ denotes the Euclidean norm, $B \in \mathbb{R}^{\alpha \times \beta}$, $y \in \mathbb{R}^\alpha$ and $\alpha, \beta \in \mathbb{N}$, $\beta \leq \alpha$. The discretization of problem 1.2.10 has not led to an overdetermined system but problem 1.3.4 can be brought though into the form of a least squares problem. The unknown quantities (V_k and ϑ_{k+1}^n, $n = 1, \ldots, N$) are linear combinations of a finite number of basis functions.

Algorithm 2 Hedged Monte Carlo

Input: u_0 ▷ seed value for random number generator

1: **if** $(N+1)\max\limits_{0\leq k<K} D_k > I$ **then**
2: break
3: **end if**
4: $(F^{(1)},\ldots,F^{(I)})=\texttt{rand}(u_0)$ ▷ generate realizations of F
5: set $V_K^{\texttt{hmc}}(\cdot)=H(\cdot)$
6: **for** $k=K-1$ to 0 **do**
7: **for** $i=1$ to I **do**
8: set $y_i = V_{k+1}^{\texttt{hmc}}(f_{k+1}(F_{0\ldots k+1}^{(i)}))$
9: **end for**
10: **for** $i=1$ to I **do**
11: set $B_{i,d}^V = b_{k,d}(f_k(F_{0\ldots k}^{(i)}))$, $d=1,\ldots,D_k$
12: set $B_{i,\cdot}^{\vartheta_n} = B_{i,\cdot}^V \cdot \Delta X_{n,k+1}(F_{0\ldots k+1}^{(i)})$, $n=1,\ldots,N$
13: **end for**
14: set $B=\left(B^V B^{\vartheta_1}\cdots B^{\vartheta_N}\right)$
15: **if** $\operatorname{rank}(B)<(N+1)D_k$ **then**
16: break
17: **else**
18: solve (1.4.1) with B and y as defined and denote the solution by x^*
19: **end if**
20: set $V_k^{\texttt{hmc}} = \sum\limits_{d=1}^{D_k} x_d^* b_{k,d}$
21: set $\vartheta_{n,k+1}^{\texttt{hmc}} = \sum\limits_{d=n\,D_k+1}^{(n+1)\,D_k} x_d^* b_{k,d}, \quad n=1,\ldots,N$
22: set $\eta_{k+1}^{\texttt{hmc}} = V_{k+1}^{\texttt{hmc}} - \vartheta_{k+1}^{\texttt{hmc}}\cdot X_{k+1}$
23: **end for**
24: set $\vartheta_0^{\texttt{hmc}}=0$ and $\eta_0^{\texttt{hmc}}=V_0^{\texttt{hmc}}$

Output: $(V_k^{\texttt{hmc}},\vartheta_k^{\texttt{hmc}},\eta_k^{\texttt{hmc}})_{k=0}^K$

Remark 1.4.1. i) `rand` is pseudo-random number generator and u_0 denotes a seed value. Concerning sampling procedures of random variables we refer to [KW08] and [Dev86].
ii) If B has not full rank, the algorithm breaks down (line 16). This occurs if the matrix B^V has not full rank or the increment of the price of two hedging instruments are linearly dependent in the sense that

$$P(\Delta X_k^{n_1} = \alpha\Delta X_k^{n_2}) = 1$$

for some n_1, n_2, $n_1\neq n_2$, k and $\alpha\in\mathbb{R}$. In the first case one may try another choice of the input parameter.

1.5. Error analysis

In §1.3 we called problem 1.3.4 a discretization of problem 1.2.10. Actually, the terms *discretization* and *approximation* support the intuition but they are somewhat meaningless since we have no convergence analysis. We analyze the error in this section.

Sources of error

We formally write

$$(V^{\texttt{lrm}}, \vartheta^{\texttt{lrm}}) := \texttt{lrm}() \qquad \text{and} \qquad (V^{\texttt{hmc}}, \vartheta^{\texttt{hmc}}) := \texttt{hmc}(u_0), \tag{1.5.1}$$

where `lrm` refers to algorithm 1 and `hmc` refers to algorithm 2. The parameter u_0 denotes the seed value for the pseudo-random number generator used to generate realizations of F. Apparently, the `lrm` algorithm has no parameter and the `hmc` algorithm only one. This is due to the fact that H, (X, N), K and (F, M) are parameters which specify the model and $(\{b_{k,d}\}_{d=1}^{D_k}, f_k)_{k=0}^{K-1}$ and I are discretization parameters and hence not parameters of the algorithms. The basic difference between the `lrm` algorithm and the `hmc` algorithm is that the `lrm` algorithm is deterministic and the `hmc` algorithm is stochastic. The output of the `hmc` algorithm depends on the I realizations of F.

The discrepancy between the output of the `hmc` and the `lrm` algorithm can be traced back to numerical error (round off and approximation) and to discretization error. The discretization error is formed by

- approximation error
  - if there is a k, $0 \leq k \leq K-1$, and a $Z \in \bigcup_{n=0}^{N} \mathcal{A}_{n,k}$ such that $\nexists\ z \in A_k$ with $Z = z \circ f_k$.
- consistency error
  - if the approximation is non-conforming in the sense that there is a $z \in A_k$ and a $n \in \{0, \ldots, N\}$ such that $\nexists\ Z \in \mathcal{A}_{n,k}$ with $Z = z \circ f_k$ for some k, $0 \leq k \leq K-1$.
- model reduction error
  - if f_k, $k = 0, \ldots, K-1$, is chosen such that information is omitted if for instance $f_k(F_{0 \ldots k}) := F_k$ is a non-Markovian model.
- stochastic error
  - if the realizations $F^{(1)}, \ldots, F^{(I)}$ do not replicate F under P exactly.

Tools for the error analysis

The error analysis presented in chapter 2 and 3 will focus on the approximation, consistency and stochastic error. The numerical error is assumed to be negligible and the models will be Markovian.

The trading strategy obtained by the hedged Monte Carlo algorithm is optimal with respect to $F^{(1)}, \dots, F^{(I)}$ but only sub-optimal with respect to F. The strategy has been determined using implicitly information about the future evolution of F.

Generating another set of realizations $F^{(I+1)}, \dots, F^{(2I)}$ of F we will compute realizations of the local and the global cost increments

$$\Delta C_{k+1}^{\mathtt{hmc}} = V_{k+1}^{\mathtt{hmc}} - V_k^{\mathtt{hmc}} - \vartheta_{k+1}^{\mathtt{hmc}} \cdot \Delta X_{k+1}, \qquad k = 0, \dots, K-1, \tag{1.5.2a}$$

$$\Delta C^{\mathtt{hmc}} := \sum_{k=0}^{K-1} \Delta C_{k+1}^{\mathtt{hmc}}. \tag{1.5.2b}$$

Note,

$$\Delta C^{\mathtt{hmc}} = V_K^{\mathtt{hmc}} - V_0^{\mathtt{hmc}} - \sum_{k=0}^{K-1} \vartheta_{k+1}^{\mathtt{hmc}} \cdot \Delta X_{k+1}.$$

The risk process $(R_k)_{k=0}^{K-1}$ is not appropriate to compare the output of the `hmc` algorithm. The expected risk R_k at time k which will arise between time k and time $k+1$ is an $\mathcal{F}_k$-measurable random variable and hence not observable[5] at time 0. We introduce estimators for the local and the global risk.

Definition 1.5.1. Let

$$\mathcal{R}_k := \left(\mathrm{Mean}\left((\Delta C_{k+1})^2\right)\right)^{1/2}, \qquad k = 0, \dots, K-1, \tag{1.5.3a}$$

$$\mathcal{R} := \left(\mathrm{Mean}\left((\Delta C)^2\right)\right)^{1/2}, \tag{1.5.3b}$$

where Mean$(\cdot)$ is defined in A.2.1a.

Remark 1.5.2. i) The risk estimator $\mathcal{R}_k$, $0 \le k \le K-1$, is closely related to R_k as the following calculations show:

$$\begin{aligned}\left(E\left[R_k \middle| \mathcal{F}_0\right]\right)^{1/2} &= \left(E\left[E\left[\Delta C_{k+1}^2 \middle| \mathcal{F}_k\right] \middle| \mathcal{F}_0\right]\right)^{1/2} \\ &= \left(E\left[\Delta C_{k+1}^2\right]\right)^{1/2} \\ &= \|\Delta C_{k+1}\|_{L^2(P)}.\end{aligned}$$

Hence, we can interpret $\mathcal{R}_k$ as an estimator of the R_k projected to time 0.
ii) If the local cost increments ΔC_{k+1}, $k = 0, \dots, K-1$, are pairwise independent,

[5] Not observable in the sense that there exists $A \in \mathcal{F}_k$ such that $0 < P(R_k^{-1}(A)) < 1$.

then

$$
\begin{aligned}
E\left[\Delta C\right]^2 &= E\left[\sum_{k=0}^{K-1} \Delta C_{k+1}\right]^2 \\
&= \sum_{k=0}^{K-1} E\left[\Delta C_{k+1}^2\right] + 2 \sum_{\substack{k_1,k_2=0 \\ k_1\neq k_2}}^{K-1} E\left[\Delta C_{k_1} \Delta C_{k_2}\right] \\
&= \sum_{k=0}^{K-1} E\left[\Delta C_{k+1}^2\right] + 2 \sum_{\substack{k_1,k_2=1 \\ k_1\neq k_2}}^{K} E\left[\Delta C_{k_1}\right] E\left[\Delta C_{k_2}\right] \\
&= \sum_{k=0}^{K-1} E\left[\Delta C_{k+1}^2\right].
\end{aligned}
$$

Hence, in this case we expect

$$
\mathcal{R}^2 \;\approx\; \sum_{k=0}^{K-1} \mathcal{R}_k^2.
$$

Locally risk-minimizing strategies are mean self-financing, cf. Theorem 1.2.7, and consequently

$$
E\left[\Delta C\right] = 0.
$$

The initial portfolio value, V_0, is therefore a natural candidate for the price of the financial derivative with payoff H. Since $(V^{\mathtt{hmc}}, \vartheta^{\mathtt{hmc}})$ is only sub-optimal,

$$
E\left[\Delta C^{\mathtt{hmc}}\right]
$$

will not be exactly zero. This means if the writer sells the financial derivative with payoff H for $V_0^{\mathtt{hmc}}$, then he has to expect to make a small loss or gain if he builds a hedge portfolio according to $(V^{\mathtt{hmc}}, \vartheta^{\mathtt{hmc}})$. To have a candidate for the price of the financial derivative which provides *on average* no loss and no gain we make the following definition.

Definition 1.5.3. Let

$$
\mathcal{V}_0 \;:=\; \text{Mean}\left(V_K - \sum_{k=0}^{K-1} \vartheta_{k+1} \cdot \Delta X_{k+1}\right). \tag{1.5.4}
$$

2. Hedging with vanilla put options

Abstract: The problem studied in this chapter is to optimal hedge a financial derivative with its underlying asset and with vanilla puts. The problem is modeled in the framework of local risk minimization. The focus is on the discretization of the problem by the hedged Monte Carlo method. We use the hedged Monte Carlo algorithm to calibrate discretization parameters by providing numerical results.

Contents

2.1. Introduction

The problem studied in this chapter is to hedge a European put H with underlying asset S, maturity K and strike R. The two hedging instruments which are available at hedging period $k \to k+1$ are the underlying S and vanilla puts Y^n with strike R_n. The hedging objective is like the second hedging instrument (Y^n) a vanilla put option. This simplifies the error analysis.
In §2.2 the problem is formulated in the framework of local risk minimization. The price of the underlying asset is modeled by a discretized geometric Brownian motion. The volatility and the drift are set constant. Hence, there is only a single source of uncertainty. The hedging instruments are therefore both used to hedge price risk. This will lead to reduction in delta and gamma exposure. In contrast to hedging based on the Black-Scholes delta and gamma, the hedging strategy is optimized in our case. We refer to [Wil07] for an introduction to hedging with Greeks.
We have conducted a small survey on the literature about optimized delta-gamma hedging. Table 2.1.1 shows the results.

Search expressions	Hits
price hedge	54700
optimal price hedge	4
delta-gamma hedge	5510
optimal delta-gamma hedge	0

Table 2.1.1.: Google search on July 19, 2011. The number of hits has been alike when we replaced *optimal* by *optimized* and/or *hedge* by *hedging*.

Hedging delta[1] exposure can be done either explicitly or implicitly. Dynamic hedging strategies are mostly explicit. The strategies are chosen such that delta exposure is reduced. Static hedging strategies on the contrary are chosen with the aim to match the payoff of the target option, cf. e.g. [BC94, DEK95, Fin03]. In this case delta is hedged implicitly. For literature on optimized static hedging we refer to [Mar09a] and the references therein.

We continue with the outline of the chapter. Section 2.3 is on the discretization of the problem formulated in §2.2 by the hedged Monte Carlo method. The error (the sum of the approximation, consistency and the stochastic error) can be large if the basis $\left(\{b_{k,d}\}_{d=1}^{D_k}\right)$ is chosen inappropriately. In §2.3 the focus is therefore on the calibration of $\{b_{k,d}\}_{d=1}^{D_k}$. We provide numerical results which have been generated for this purpose. Furthermore §2.3 contains a lot of practical information concerning the application of the hedged Monte Carlo method and the implementation of the hedged Monte Carlo algorithm.

[1] Or any other Greek.

2.2. Problem formulation

Local risk minimization amounts to define the hedging objective H, the number of hedging opportunities K, the number and the price of the hedging instruments (N, X) and an associated filtered probability space $(\Omega, \mathcal{F}, P, (\mathcal{F}_k)_{k=0}^K)$.

The hedging objective is a European put option H with strike price R, $R \in \mathbb{R}_+$, maturity K, $K \in \mathbb{N}$, and underlying asset S. The payoff function of H is

$$H(x) = \max(R - x, 0), \qquad x > 0.$$

The price of the underlying at time k is

$$S_{k+1} := S_k \left(1 + \mu\delta t + \sigma\sqrt{\delta t} Z_{k+1}\right), \quad k = 0, \dots, K-1, \tag{2.2.1a}$$

$$S_0 := s_0 \tag{2.2.1b}$$

with $s_0 \in \mathbb{R}_+$, $\delta t := \frac{T}{K}$, $T \in \mathbb{R}_+$ and $(Z_k)_{k=1}^K$ are i.i.d. random variables with $Z_1 \sim N(0,1)$. The drift $\mu \in \mathbb{R}_+$ and the volatility $\sigma \in \mathbb{R}_+$ are constant and $\mathcal{F}_0$-measurable, i.e. known at time 0. The price of the numéraire asset at time k is

$$B_k := \exp(rk\delta t) \quad k = 0, \dots, K \tag{2.2.2}$$

with constant $\mathcal{F}_0$-measurable interest rate $r \in \mathbb{R}_+$. We set

$$X_k^1 := S_k / B_k, \quad k = 0, \dots, K.$$

The price of N_Y, $N_Y := K$, European put options Y^n, $1 \le n \le N_Y$, with maturity K, strike R_n, underlying S and payoff function

$$Y^n(x) = \max(R_n - x, 0), \qquad x > 0$$

is at time k, $k = 0, \dots, K$, given by

$$Y_k^n := \texttt{bs\_put\_price}(S_k, T - k\delta t; R_n, r, \sigma) \quad n = 1, \dots, N_Y, \tag{2.2.3}$$

where `bs_put_price` is the Black-Scholes formula (see A.3.1). Finally, we set

$$X_k^{1+n} := Y_k^n / B_k, \quad k = 0, \dots, K$$

for $n = 1, \dots, N_Y$.

Remark 2.2.1. i) The model defined above is similar to the Black-Scholes model. The recursive scheme (2.2.1) is sometimes called Euler-Maruyama scheme or stochastic Euler scheme. It allows to compute a first order approximation to the solution of the stochastic differential equation

$$dS_t = S_t (\mu dt + \sigma dW_t), \quad t > 0 \tag{2.2.4}$$

with initial condition $S_0 = s_0$ and Wiener process $(W_t)_{t\geq 0}$. Weak and strong convergence are established for instance in [KP92]. In [BP03] and [BS94] it is shown that if the price of the option Y^n, $1 \leq n \leq N_Y$, is defined by its hedging costs based on a strategy obtained by local risk minimization, then the price converges to its Black-Scholes price.
ii) If $(S_k, B_k)_{k=0}^K$ is chosen in some other way, one may analyze whether definition (2.2.3) is still reasonable.
iii) Using the Black-Scholes formula is of computational advantage, since its evaluation can be done very fast.

While the markets in the underlying and in the numéraire asset are liquid at any time instance k, a position in Y^n can only be constructed at time $k = n - 1$. At time $k + 1$ this position has to be liquidated. That means there are exactly two hedging instruments available on each hedging period $k \to k+1$. In the framework of local risk minimization liquidity is no issue. All assets are liquid. If the liquidity has to be limited, one has to redefine the spaces of admissible trading strategies $\mathcal{A}^r_{n,k}$, $n = 0, \ldots, N$. If an asset is not traded at all at some time instance k, one sets $\mathcal{A}^r_{n,k} := \{\, Z : \Omega \to \mathbb{R} \,|\, Z = 0 \;\; P\text{-a.s.} \,\}$. For the scenario described above we (formally) set

$$\mathcal{A}^r_{1+n,k} := \{\, Z : \Omega \to \mathbb{R} \,|\, Z = 0 \;\; P\text{-a.s.} \,\}, \quad k = 0, \ldots, n-2, n, \ldots, K-1$$

for $n = 1, \ldots, N_Y$.
The problem is then to find a locally risk-minimizing strategy (η, ϑ).

Convention 2.2.2. For notational simplicity we will denote the second hedge, $(\vartheta^2, \ldots, \vartheta^{1+N_Y})$, by λ, i.e. $\lambda_0 = 0$ and

$$\lambda_k := \vartheta_k^{k+1}, \quad k = 1, \ldots, K.$$

The first hedge ϑ^1 will be denoted by ϑ if it is clear from the context.

The numerical results presented below have been obtained using the following parameters:

$$\mu = 0.05,\ \sigma = 0.3,\ s^0 = 100,\ r = 0.05$$

and

$$T = 1,\ R = 100,\ R_n = S_{k=n-1} \text{ for } n = 1, \ldots, N_Y$$

as well as

$$K = 20.$$

Note, $Y^n \neq H$ for $n = 2, \ldots, N_Y$ but $Y^1 \equiv H$. This however is no problem. The framework of §1.2 includes the case $Y^1 \equiv H$. Actually, it has been a nice opportunity to check our computer code. The optimal hedge from time 0 to time 1 obtained by the hedged Monte Carlo method has to be 'hold no stocks S and one option Y^1', i.e. the perfect hedge which provides no residual risk from time 0 to time 1. Since the algorithm runs backward in time, the perfect hedge on $0 \to 1$ does not improve or worsen the hedges at later time instances.

2.3. Problem discretization by the hedged Monte Carlo method

The discretization of (1.2.10) according to the hedged Monte Carlo methods amounts to specify the $\mathbb{R}^M$-valued process F which generates the filtration $(\mathcal{F}_k)_{k=0}^K$, and the discretization parameters $\left(\{b_{k,d}\}_{d=1}^{D_k}, f_k\right)_{k=0}^{K-1}$ and I, cf. §1.3.
The application of the hedged Monte Carlo method is easy. The model tells us how to choose F and $(f_k)_{k=0}^K$, cf. §2.3.1. We have observed that independently of the choice of $\left(\{b_{k,d}\}_{d=1}^{D_k}\right)_{k=0}^{K-1}$ increasing I leads to smaller variance of the output of the hedged Monte Carlo algorithm. Hence, the choice of I seems to be a matter of the allowed variance and the available computing time. We set I to a moderate value in §2.3.1. Choosing $\left(\{b_{k,d}\}_{d=1}^{D_k}\right)_{k=0}^{K-1}$ is most demanding. Simply increasing D_k does not necessarily improve the approximation. In §2.3.3 we show how a good choice of $\left(\{b_{k,d}\}_{d=1}^{D_k}\right)_{k=0}^{K-1}$ can be determined and what a good choice of $\left(\{b_{k,d}\}_{d=1}^{D_k}\right)_{k=0}^{K-1}$ actually is.

2.3.1. The calibration framework

We have set $M = 1$, $F_k := S_k$, $k = 0, \ldots, K$ and $I = 5\,000$. The functions $f_k : F_{0\ldots k}(\Omega) \subset \mathbb{R}^{(k+1)M} \to U_k$, $k = 0, \ldots, K-1$ are defined next. Afterwards, we define the class of (basis) functions $b_{k,d} : U_k \to \mathbb{R}$ which constitutes the framework for the calibration in §2.3.3.

a) Since $(S_k)_{k=0}^K$ generates the filtration $(\mathcal{F}_k)_{k=0}^K$ and since S is a Markov process, we set

$$f_k(S_0, \ldots, S_k) := S_k, \quad k = 0, \ldots, K-1.$$

We further set

$$U_k = \begin{cases} \{S^0\} & \text{if } k = 0 \\ (0, \infty) & \text{if } 1 \le k \le K-1. \end{cases}$$

Note that the recursion (2.2.1) does not prohibit $S_k \le 0$. The probability of the event

$$\tilde{\Omega}_k := \left\{ Z_k < \frac{-1 - \mu\delta t}{\sigma\sqrt{\delta t}} \right\} \subset \Omega$$

is in general very small; for the chosen parameters

$$P\left(\tilde{\Omega}_k\right) = 8.5 \times 10^{-51}.$$

We have never observed this event.

Remark 2.3.1. An alternative to (2.2.1a) is to define $(S_k)_{k=1}^K$ by

$$S_{k+1} := S_0 \exp\left(\left(\mu - \frac{\sigma^2}{2}\right)(k+1)\delta t + \sigma\sqrt{(k+1)\delta t} Z_{k+1}\right), \quad k = 0, \ldots, K-1.$$

In this case S_k is always positive.

b) We define the class of bases $\{b_{k,d}\}_{d=1}^{D_k}$. Since $\operatorname{card}(U_0) = 1$, we set $D_0 = 1$ and $b_{0,1} := 1$. For $0 < k < K$ we set $D_k = D$ where $D \in \mathbb{N}$ and

$$b_{k,d} = \begin{cases} L_d^w & \text{or} \\ N_{d,\ell} & \end{cases}, \quad d = 1, \ldots, D,$$

where L_d^w is the d-th weighted Laguerre polynomial (see appendix A.1.1 for definition) and $N_{d,\ell}$, $1 \le \ell \le 4$, is the d-th B-spline function of order ℓ (see appendix A.1.2 for definition).

Remark 2.3.2. In view of the literature on least squares Monte Carlo methods, cf. e.g. [LS98, MN01] the standard choice seems to be to take orthogonal polynomials like Laguerre polynomials. Laguerre polynomials have global support. To have something in contrast we made the decision for B-splines, i.e. functions of local support.

B-splines are defined with respect to a set of nodes. This set has been chosen adaptively. The number of required nodes depends on the B-spline order and the number of basis functions (see appendix A.1.2). Assume $n+1$ nodes $\{x_{k,0} < \ldots < x_{k,n}\}$ have to be specified. Given I realizations $S_k^{(1)}, \ldots, S_k^{(I)}$ of S_k the nodes have been set to

$$x_{k,j} = \begin{cases} \min\limits_{i=1,\ldots,I} S_k^{(i)} & \text{if } j = 0 \\ \max\limits_{i=1,\ldots,I} S_k^{(i)} & \text{if } j = n \\ \min\left\{ S_k^{(i)} \,\middle|\, \operatorname{card}\{S_k^{(m)} \,|\, S_k^{(m)} \le S_k^{(i)}\} = \lfloor I/n \rfloor \cdot j \right\} & \text{if } 0 < j < n \end{cases}$$

for $k = 1, \ldots, K-1$. We will use the following notation in Figure 2.3.1:

$$\underline{x}_k = x_{k,0} \quad \text{and} \quad \bar{x}_k = x_{k,n}.$$

2.3.2. Implementation of the hedged Monte Carlo algorithm

The least squares problem (1.4.1) has been solved by QR-factorization by means of Householder transformations. This requires about $2 \times \alpha \times \beta^2$ operations if $\alpha \gg \beta$ [Sie06].

All numerical results presented in this chapter have been obtained on a Intel(R) Core(TM)2 Duo CPU E8400 @ 3.00GHz using a single core with MHz 1998.000 and cache size 6144 KB. The computer code has been written in `C++`.

Weighted Laguerre polynomials (weight $e^{-x/2}$) rapidly approach zero for $x \to \infty$. If $x \gg 1$, this causes numerical imprecision. We have used `double` for real numbers. As a remedy we have replaced internally s_0 by s_0/s_0, etc. and afterwards transformed back.

2.3.3. Calibration of discretization parameters

The objective of this section is to calibrate the basis $\{b_{k,1}, \ldots, b_{k,D}\}$. Generally speaking we consider the basis as *best* which gives rise to minimal error (approximation + consistency + stochastic), cf. §1.5. For practical matter we further require minimal computing time of the hedged Monte Carlo algorithm.

Our tool for calibration will be the hedged Monte Carlo algorithm and our methodology will be simulation. For a given parameter specification we use the hedged Monte Carlo algorithm to compute hedging strategies and then we investigate their performance by simulation. The performance measures will be the risk estimator $\mathcal{R}$ and the estimator of the initial portfolio value $\mathcal{V}_0$ (see §1.5 for definition).
The use of the hedged Monte Carlo algorithm introduces some noise whose size depends on I. Actually, we do not know if $\mathcal{R}$ and $\mathcal{V}_0$ are biased. We assume that they are unbiased and set up the following criteria. Choose D and $b_{k,1}, \ldots, b_{k,D}$ such that

$$\left|E[\mathcal{R}^{\texttt{hmc}}] - E[\mathcal{R}^{\texttt{lrm}}]\right| \to \min \tag{2.3.1a}$$
$$\mathrm{Var}[\mathcal{R}^{\texttt{hmc}}] \to \min \tag{2.3.1b}$$

and

$$\left|E[\mathcal{V}_0^{\texttt{hmc}}] - E[\mathcal{V}_0^{\texttt{lrm}}]\right| \to \min \tag{2.3.1c}$$
$$\mathrm{Var}[\mathcal{V}_0^{\texttt{hmc}}] \to \min \tag{2.3.1d}$$

as well as

$$\tau^{\texttt{hmc}} \to \min \tag{2.3.1e}$$

where $\tau^{\texttt{hmc}}$ denotes the computing time for $(V^{\texttt{hmc}}, \vartheta^{\texttt{hmc}}, \lambda^{\texttt{hmc}})$. Since we do not know $(V^{\texttt{lrm}}, \vartheta^{\texttt{lrm}}, \lambda^{\texttt{lrm}})$, we replace $(V^{\texttt{lrm}}, \vartheta^{\texttt{lrm}}, \lambda^{\texttt{lrm}})$ by the Black-Scholes hedging strategy $(V^{\mathsf{bs}}, \vartheta^{\mathsf{bs}}, \lambda^{\mathsf{bs}})$ defined below. The strategy is close to optimal, cf. Remark 2.2.1.

Remark 2.3.3. It was not our objective to give a proper mathematical formulation of the problem of choosing D and $b_{k,1}, \ldots, b_{k,D}$. With the five criteria we just want to show a way how one can find a good basis. The reader may decide which basis is *best* (in his eyes).

The Black-Scholes hedging strategy

Let $H_k^{\mathsf{bs}}(S) := \texttt{bs\_put\_price}(S, T - k\delta t; R, r, \sigma)$, $k = 0, \ldots, K-1$, cf. (A.3.1), and let

$$\Pi_{n,k}^{\mathsf{bs}}(S, \vartheta, \lambda, \eta) := -H_k^{\mathsf{bs}}(S) + \vartheta S + \lambda Y_{n,k}(S) + \eta.$$

The interpretation is that $\Pi_{n,k}^{\mathsf{bs}}$ describes the portfolio value of a portfolio consisting of -1 options H, ϑ shares of the underlying S, λ shares of the options Y_n and η shares of the numéraire asset.

The Black-Scholes hedging strategy $(\vartheta^{\mathsf{bs}}_{k+1}, \lambda^{\mathsf{bs}}_{k+1})$, $k = 0, \dots, K-1$, is defined as the solution to

$$\frac{\partial}{\partial S}\Pi^{\mathsf{bs}}_{n,k}(S, \vartheta^{\mathsf{bs}}_{k+1}, \lambda^{\mathsf{bs}}_{k+1}, \eta) = 0 \tag{2.3.2a}$$

$$\frac{\partial^2}{\partial S^2}\Pi^{\mathsf{bs}}_{n,k}(S, \vartheta^{\mathsf{bs}}_{k+1}, \lambda^{\mathsf{bs}}_{k+1}, \eta) = 0. \tag{2.3.2b}$$

For $k = 0, \dots, K-1$ this gives

$$\lambda^{\mathsf{bs}}_{k+1}(x) := \frac{\Gamma^{\mathsf{bs}}_k(x, R)}{\Gamma^{\mathsf{bs}}_k(x, x)}, \tag{2.3.3}$$

$$\vartheta^{\mathsf{bs}}_{k+1}(x) := \Delta^{\mathsf{bs}}_k(x, s_0) - \lambda^{\mathsf{bs}}_{k+1}(x) \cdot \Delta^{\mathsf{bs}}_k(x, x) \tag{2.3.4}$$

with

$$\begin{aligned} \Delta^{\mathsf{bs}}_k(S; K) &:= \texttt{bs\_put\_delta}(S, T - k\delta t; K, r, \sigma), \\ \Gamma^{\mathsf{bs}}_k(S; K) &:= \texttt{bs\_gamma}(S, T - k\delta t; K, r, \sigma), \end{aligned}$$

cf. (A.3.2) and (A.3.3). The portfolio value $(V^{\mathsf{bs}}_k)_{k=0}^K$ is defined in the Black-Scholes case by $V^{\mathsf{bs}}_k := H^{\mathsf{bs}}_k$, $k = 0, \dots, K$.

2.3.3.1. Numerical results

Table 2.3.1 shows the results obtained using various choices of the basis

$$A_k = \begin{cases} \operatorname{span}\{L^w_1, \dots, L^w_D\} & D = 3, \dots, 8 \\ \operatorname{span}\{N_{1,\ell}, \dots, N_{D,\ell}\} & D = (3,)4, 6, 8, 10, 15, 20, \ \ell = 1, \dots, 4. \end{cases}$$

In view of Table 2.3.1, we draw the following conclusions:

Laguerre polynomials do not seem to be appropriate: $D = 4$ has been the best choice. However, the mean of $\mathcal{R}$ and the standard deviation of $\mathcal{V}_0$ and $\mathcal{R}$ have been much larger than in the case using the Black-Scholes strategy. The results turned worse with increasing D since the polynomials start to oscillate strongly.

Splines of order 1 are piecewise constant functions and hence have not shown this phenomenon. The results became better with increasing D but about 20 basis functions have to be used in order to get results of the same quality as in the case of 4 Laguerre polynomials. If $D = 20$, the computing time is however about 10-times larger than in the case $D = 4$.

The results turned much better with splines of order 2-4. With $D \geq 8$ we could establish results of about the same quality as those obtained by the Black-Scholes strategy. Among them, a particular good choice is using 10 B-splines of order 3: the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator seem to be unbiased with respect to Black-Scholes and have about the same standard deviation. The optimal strategy has been obtained in 1.3 seconds.

2.3.3.2. Visualization of a hedging strategy

In this section we provide some illustration of our findings in the last section. To that end we have plotted for two different time instances (k=1,19) the portfolio value and the first and the second hedge as a function of the price of the underlying, see Figure 2.3.1. In each of the six cases we compare the curves obtained using 10 resp. 20 B-splines of order 3 resp. 4 Laguerre polynomials and the Black-Scholes counterpart.

In view of the V_1-regression (subfigure (a)) it is hard to recognize a difference between the three bases and the Black-Scholes curve. The V_{19}-regression (subfigure (b)) already indicates that Laguerre polynomials are not capable to reproduce curves with fast changing curvature. This becomes even more evident at the ϑ_{20}- and λ_{20}-regression (subfigures (d) and (f)).
The first plot where a difference between the B-splines bases and the Black-Scholes curve is visable is the plot which shows the ϑ_2-regression (subfigure (c)). The curves do not coincide where the realizations lie less dense (cf. as well subfigure (d), (e) and (f)).
The regression of λ_2 (subfigure (e)) shows what can happen when the number of basis functions gets too large. The curve for 20 B-splines has a small cap between $x = 90$ and $x = 95$. The little cap around $x = 105$ in subfigure (d) results from the fact that there are two options which have to be hedged by the underlying asset (cf. formula (2.3.4)).

In view of Figure 2.3.1 we can observe two phenomena which take place independently of the chosen basis.
The regression seems to be more difficult close to maturity. In the Black-Scholes model there is an easy explanation for this phenomena. The option price is the solution of a parabolic partial differential equation (pde), the Black-Scholes (pde) with final time condition the payoff function. The payoff function is not differentiable at the strike price. However, the solution of a parabolic pde smoothes out with passage of time and smooth curves are easier to approximate than not so smooth curves.
The second observation is that the portfolio value is better regressed than the first hedge which is in turn better regressed than the second hedge. For this phenomena there is as well an explanation in the Black-Scholes model. The hedging strategy $(\vartheta^{\mathsf{bs}}, \lambda^{\mathsf{bs}})$ as defined in (2.3.4) and (2.3.3) has more complicated shapes than the price of H.

The last two observations show that it is not optimal to have constant $\dim(A_k)$, $k > 0$. To reduce the computing time one may choose $D_k := \dim(A_k)$, $k > 0$, such that D_k is decreasing with decreasing k. Besides, one may choose three different spaces A_k^V, A_k^ϑ and A_k^λ for V_k, ϑ_{k+1} and λ_{k+1} with $\dim(A_k^V) \leq \dim(A_k^\vartheta) \leq \dim(A_k^\lambda)$.

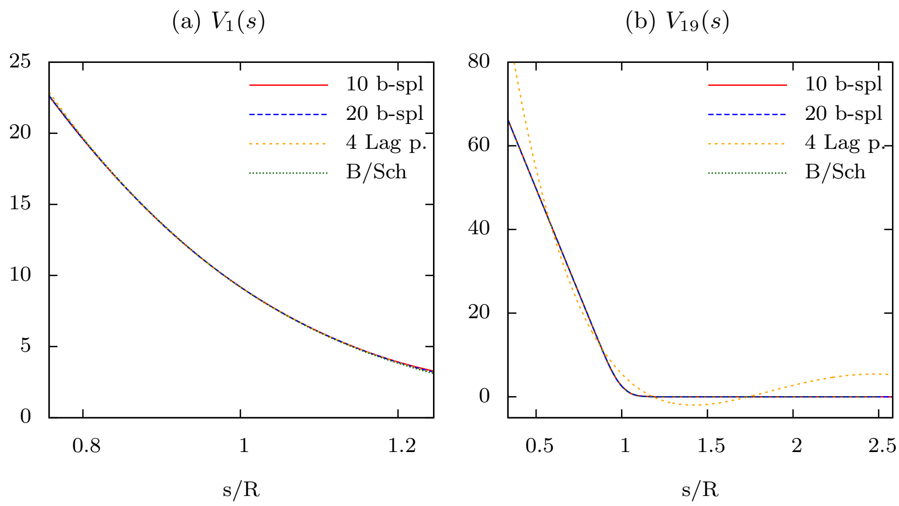


Figure 2.3.1.: The plots show the optimal hedging strategy obtained by least squares regression using the hedged Monte Carlo algorithm for three different choices of the basis of A_k, $k > 0$. The bases have been 10 resp. 20 B-splines of order 3 resp. the first 4 Laguerre polynomials. The curve called 'B/Sch' is $H^{\mathsf{bs}}(s, k\delta t)$ (in subfigures (a) and (b)),$\vartheta^{\mathsf{bs}}_{k+1}(s)$ (in subfigures (c) and (d)), and $\lambda^{\mathsf{bs}}_{k+1}(s)$ (in subfigures (e) and (f)). The x-range of the plots is $[\underline{x}_k, \bar{x}_k]$ cf. §2.3.2

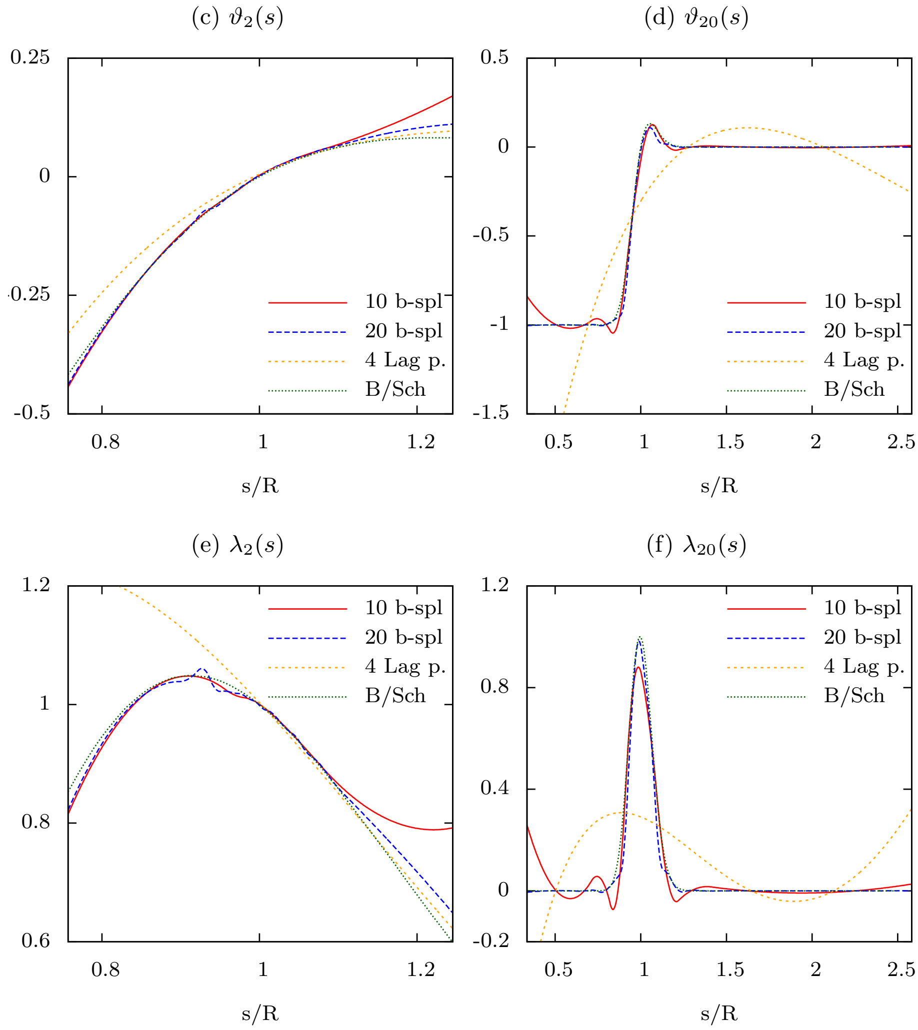


Figure 2.3.1.: The plots show the optimal hedging strategy obtained by least squares regression using the hedged Monte Carlo algorithm for three different choices of the basis of A_k, $k > 0$. The bases have been 10 resp. 20 B-splines of order 3 resp. the first 4 Laguerre polynomials. The curve called 'B/Sch' is $H^{\mathsf{bs}}(s, k\delta t)$ (in subfigures (a) and (b)),$\vartheta^{\mathsf{bs}}_{k+1}(s)$ (in subfigures (c) and (d)), and $\lambda^{\mathsf{bs}}_{k+1}(s)$ (in subfigures (e) and (f)). The x-range of the plots is $[\underline{x}_k, \bar{x}_k]$ cf. §2.3.2

(a) Laguerre polynomials

D	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	
3	-0.0074	5.3467	5.2726	4.3694	0.17
4	-0.0041	4.5333	4.5068	4.2117	0.23
5	0.0025	5.4267	5.0103	44.1486	0.29
6	-0.0057	6.8000	5.1701	95.6441	0.36
7	-0.0025	15.0400	7.2840	293.3243	0.45
8	0.1807	377.0267	44.4020	7894.5856	0.55

(b) b-splines of order 1

D	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	
3	0.0223	16.5733	16.3631	6.5405	0.21
4	0.0202	13.6800	13.7356	6.9910	0.27
6	0.0144	10.1467	10.4409	6.9820	0.39
8	0.0103	8.4000	8.5502	6.8514	0.56
10	0.0090	7.0267	7.3421	6.7928	0.79
15	0.0056	5.4533	5.6024	6.4369	1.55
20	0.0031	4.4933	4.6711	6.1577	2.61

(c) b-splines of order 2

D	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	
3	0.0041	5.6400	5.7745	2.8514	0.24
4	0.0025	3.0400	2.8552	1.4324	0.31
6	0.0001	1.6400	1.5564	0.7928	0.45
8	0.0002	1.2667	1.2210	0.7793	0.64
10	0.0003	1.1600	1.1017	0.8784	0.89
15	-0.0002	1.0933	1.0271	0.9775	1.70
20	-0.0001	1.0533	1.0152	1.0270	2.81

Table 2.3.1.: Sample mean (m) and sample standard deviation (s) of the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator using 250 samples (column 2-5); the superscripts (hmc,bs) refer to the underlying hedging strategy of the $\mathcal{V}_0$- and $\mathcal{R}$-estimator; the results obtained by the Black-Scholes strategy ($\vartheta^{\texttt{bs}}, \lambda^{\texttt{bs}}$) have been: $\mathrm{m}^{\texttt{bs}}(\mathcal{V}_0) = 9.3571$, $\mathrm{s}^{\texttt{bs}}(\mathcal{V}_0) = 0.0075$, $\mathrm{m}^{\texttt{bs}}(\mathcal{R}) = 0.5525$, $\mathrm{s}^{\texttt{bs}}(\mathcal{R}) = 0.0222$; $\tau_D^{\texttt{hmc}}$ denotes the time required to compute ($V^{\texttt{hmc}}, \vartheta^{\texttt{hmc}}, \lambda^{\texttt{hmc}}$); $\tau^r = 1.32$ sec.; $D = \dim(A_k)$ for $k > 0$ the number of basis functions.

(d) b-splines of order 3

D	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	
3	-0.0109	8.2000	7.9517	15.4730	0.27
4	-0.0065	6.0267	5.8206	7.6081	0.36
6	0.0002	1.7600	1.7081	3.1126	0.52
8	0.0001	1.1867	1.1493	1.0450	0.73
10	0.0000	1.0533	1.0340	0.9459	1.00
15	-0.0002	1.0267	0.9975	1.0045	1.86
20	-0.0002	1.0133	0.9991	1.0270	3.03

(e) b-splines of order 4

D	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^{\texttt{bs}}$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^{\texttt{bs}}$	
4	-0.0001	4.7467	4.6375	11.6802	0.39
6	0.0022	3.6267	3.3723	19.1577	0.60
8	0.0007	1.6400	1.4123	7.9505	0.83
10	0.0001	1.1067	1.0713	1.7072	1.11
15	0.0000	1.0267	0.9966	1.0721	2.03
20	-0.0001	1.0267	0.9978	1.0405	3.24

Table 2.3.1.: Sample mean (m) and sample standard deviation (s) of the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator using 250 samples (column 2-5); the superscripts (`hmc`,`bs`) refer to the underlying hedging strategy of the $\mathcal{V}_0$- and $\mathcal{R}$-estimator; the results obtained by the Black-Scholes strategy ($\vartheta^{\texttt{bs}}, \lambda^{\texttt{bs}}$) have been: $\mathrm{m}^{\texttt{bs}}(\mathcal{V}_0) = 9.3571$, $\mathrm{s}^{\texttt{bs}}(\mathcal{V}_0) = 0.0075$, $\mathrm{m}^{\texttt{bs}}(\mathcal{R}) = 0.5525$, $\mathrm{s}^{\texttt{bs}}(\mathcal{R}) = 0.0222$; $\tau_D^{\texttt{hmc}}$ denotes the time required to compute ($V^{\texttt{hmc}}, \vartheta^{\texttt{hmc}}, \lambda^{\texttt{hmc}}$); $\tau^r = 1.32$ sec.; $D = \dim(A_k)$ for $k > 0$ the number of basis functions.

3. Hedging with variance swaps

Abstract: This chapter is about optimal hedging price and volatility exposure of a European put option. The hedging instruments are the underlying asset and variance swaps. The problem is modeled in the framework of local risk minimization in discrete time. We focus on the appropriate discretization of the problem by the hedged Monte Carlo method. We investigate the distribution of the hedging costs and the residual risk. We address the efficiency of the hedged Monte Carlo algorithm. Numerical results are provided and constitute the basis of the analysis.

Contents

3.1. Introduction

An introduction to volatility derivatives in general and to variance and volatility swaps/ futures in particular can be found in [DDKZ99, BSG05, BW07]. An excellent survey on volatility derivative is [CL09].

Traders use variance swaps to hedge volatility exposure. These instruments provide pure exposure to volatility and hence do not require the hedging of directional risk. Directional risk arises if options are used instead of variance swaps to hedge volatility exposure. The price of options depends strongly on the price of their underlying asset and not only on the volatility.

The hedge problem of this chapter features a European put option as hedging objective, two hedging instruments and two sources of uncertainty. The hedging instruments are the underlying asset of the hedging objective and variance swaps. The variance swaps and the hedging objective share the same underlying asset. The sources of uncertainty are the price of the underlying asset and its variance/volatility. The basic problem is to find an optimal dynamic hedging strategy. In section 3.2 the problem is formulated in the framework of local risk minimization, cf. §1.2.

In view of the available hedging instruments and the present sources of uncertainty it can be deduced that the hedging strategy will be to reduce price and volatility exposure - delta and in part gamma and vega exposure. We have conducted a small survey on the literature about optimal hedging of price and volatility exposure. Table 3.1.1 shows the results.

Search expression	Hits
price and volatility hedge	1800
optimal price and volatility hedge	0
delta-gamma-vega hedge	866
optimal delta-gamma-vega hedge	0

Table 3.1.1.: Google search on July 21, 2011. The number of hits has been alike when we replaced *optimal* by *optimized* and/or *hedge* by *hedging*.

Numerical results are provided in §3.4-§3.6. Section 3.4 is on the discretization of the local risk minimization problem. We use the hedged Monte Carlo algorithm to calibrate discretization parameters. In section 3.5 hedging under local risk minimization is investigated from an economic point of view. The analysis addresses the distribution of the hedging costs and the residual risk for various hedging frequencies and one or two hedging instruments. The efficiency of the hedged Monte Carlo algorithm is finally investigated in section 3.6.

3.2. Problem formulation

Local risk minimization amounts to define the hedging objective H, the number of hedging opportunities K, the number and the price of the hedging instruments (N, X) and an associated filtered probability space $\left(\Omega, \mathcal{F}, P, (\mathcal{F}_k)_{k=0}^K\right)$, cf. §1.2.

The hedging objective is a European put option with strike R, $R \in \mathbb{R}_+$, maturity K, $K \in \mathbb{N}$, underlying asset S and payoff function

$$H(x) = \max(R - x, 0), \qquad x > 0.$$

The price S_k of the underlying at time k, $k = 0, \dots, K$, is such that

$$S_{k+1} = S_k \left(1 + \mu\delta t + \sigma_k\sqrt{\delta t}Z_k^1, \right), \quad k = 0, \dots, K-1 \tag{3.2.1a}$$

$$S_0 = s_0 \tag{3.2.1b}$$

with $s_0 \in \mathbb{R}_+$, $\delta t := \frac{T}{K}$, $T \in \mathbb{R}_+$, and drift $\mu \in \mathbb{R}_+$. The volatility $(\sigma_k)_{k=0}^K$ is such that

$$\sigma_k = \sqrt{\nu_k}, \quad k = 0, \dots, K \tag{3.2.2a}$$

and the variance $(\nu_k)_{k=0}^K$ is such that

$$\nu_{k+1} = \nu_k \left(1 + f\sqrt{\delta t}Z_k^2\right), \quad k = 0, \dots, K-1 \tag{3.2.2b}$$

$$\nu_0 = \nu^0 \tag{3.2.2c}$$

with initial variance ν^0, $\nu^0 \in \mathbb{R}_+$, and volatility of the variance f, $f \in \mathbb{R}_+$. It is assumed that $(Z_k^\ell)_{k=1}^K$ are i.i.d. random variables with $Z_k^\ell \sim N(0,1)$ for $\ell = 1, 2$ and that Z_k^1 and Z_k^2, $k = 0, \dots, K-1$, have constant correlation ρ, $|\rho| \in (0,1)$, i.e.

$$\mathrm{Cov}(Z_k^1, Z_k^2) = \rho, \quad k = 0, \dots, K-1.$$

There are numéraire assets B with price $B_k := \exp(rk\delta t)$, $k = 0, \dots, K$, and variance swaps Y with price Y_k at time k, $k = 0, \dots, K$.

Remark 3.2.1. The definition of $(S_k)_{k=0}^K$, equation (3.2.1), shows that the volatility today (σ_k) is the future gamma (S_{k+1}).

To put this into the framework described in §1.2 we set $N = 2$ and

$$X_k^1 := S_k/B_k, \quad k = 0, \dots, K, \tag{3.2.3a}$$

$$X_k^2 := \frac{k}{K}\tilde{\nu}_k^R + \frac{K-k}{K}\tilde{\nu}_k - \nu_0, \quad k = 0, \dots, K \tag{3.2.3b}$$

where

$$\tilde{\nu}_k^R = \begin{cases} 0 & \text{if } k = 0 \\ \frac{K}{k}\sum_{\ell=1}^k \left(\log\left(\frac{X_\ell^1}{X_{\ell-1}^1}\right)\right)^2 & \text{if } k = 1, \dots, K \end{cases}$$

and

$$\tilde{\nu}_k := \nu_k / B_k, \qquad k = 0, \dots, K.$$

For completeness, $(Y_k)_{k=0}^K$ is defined by $Y_k = B_k X_k^2$, $k = 0, \dots, K$. Motivation and interpretation of (3.2.3b) can be found in the next section.

Remark 3.2.2. The process which describes the price of the variance swaps should be defined such that it is *consistent* with the price of the underlying and the variance. Furthermore the dynamics of the process shall feature a variance swap. In order to find a candidate for this process one may consider a model which is *close* to the model under work and which is complete. In a complete model with risk neutral probability measure, the price of a contingent claim is unique. For the case here, we consider the continuous time counterpart of the model, cf. section 3.3.

The problem is then to find a locally risk-minimizing strategy.

The numerical results we present below have been obtained with

$$\mu = 0.05,\ s_0 = 100,\ \rho = -0.5,\ f = 0.3,\ \nu_0 = 0.09,\ r = 0.05$$

and

$$T = 1,\ R = 100.$$

3.3. Valuation of variance swaps

The reader may skip this section. Later sections will not draw from this section.

The model. Consider the following system of stochastic differential equations

$$\begin{aligned} dB_t &= rB_t\,dt, & t>0, && B_0 &= 1, && (3.3.1a)\\ dS_t &= \mu S_t\,dt + \sqrt{\nu_t} S_t\,dW_t^1, & t>0, && S_0 &= s_0, && (3.3.1b)\\ d\nu_t &= a(\nu_t,t)\,dt + f\nu_t\,dW_t^2, & t>0, && \nu_0 &= \nu_0, && (3.3.1c)\end{aligned}$$

where $r, \mu, s_0, f, \nu_0 \in \mathbb{R}_+$ and $(W_t^\ell)_{t\geq 0}$, $\ell = 1,2$, are Wiener processes with constant correlation ρ, $|\rho| \in (0,1)$, defined on a probability space $(\Omega, \mathcal{F}, P)$ with filtration $(\mathcal{F}_t)_{t\geq 0}$, $\mathcal{F}_t \subset \mathcal{F}$, $t \geq 0$, generated by $(W_t^\ell)_{t\geq 0}$, $\ell = 1,2$ made complete and right-continuous.

The coefficient $a(\cdot,\cdot)$ is assumed to be such that the system (3.3.1) has a unique solution (B, S, ν) and such that there exists a unique martingale measure $P^\star$ under which $\tilde{S} = (\tilde{S}_t)_{t\geq 0}$, $\tilde{S}_t := S_t/B_t$, $t \geq 0$ and $\tilde{\nu}_t = (\tilde{\nu}_t)_{t\geq 0}$, $\tilde{\nu}_t := \nu_t/B_t$, $t \geq 0$ are martingales.

System (3.3.1) is sometimes called a stochastic volatility model.

Valuation of financial derivatives. Let $T \in \mathbb{R}_+$ and $\tilde{Y}_T$ a $\mathcal{F}_T$-measurable and $\tilde{P}$-integrable random variable and let $(\tilde{Y}_t)_{0\leq t\leq T}$ be a stochastic process defined by

$$\tilde{Y}_t = E^\star\left[\tilde{Y}_T | \mathcal{F}_t\right], \quad 0 \leq t \leq T, \tag{3.3.2}$$

where $E^\star$ denotes the expectation with respect to $P^\star$.

Interpretation 3.3.1. Think of $\tilde{Y}_T$ as the discounted payoff of a financial derivative Y with maturity T and European exercise feature. Assume that the model is complete and hence Y has a unique price $\tilde{Y}$. The price $\tilde{Y}$ has to be a martingale since otherwise there would be some market participants who would out-arbitrage this (no-arbitrage principle). We refer to [BK04] for an introduction to the risk-neutral valuation of financial derivatives.

Valuation of variance swaps. Let

$$\tilde{Y}_T = \tilde{\nu}_T^R - \tilde{R}_Y, \tag{3.3.3}$$

where ν_T^R is a $\mathcal{F}_T$-measurable and $P^\star$-integrable random variable and $\tilde{R}_Y$ is a $\mathcal{F}_0$-measurable random variable defined by

$$\tilde{R}_Y := E^\star \left[\tilde{\nu}_T^R | \mathcal{F}_0\right].$$

In this case the price of Y at time t is

$$\tilde{Y}_t = \left(E^\star \left[\tilde{\nu}_T^R | \mathcal{F}_t\right] - \tilde{R}_Y\right), \quad 0 \leq t \leq T$$

and in particular $\tilde{Y}_0 = 0$ $P^\star$-a.s.

Interpretation 3.3.2. The financial derivative Y is a variance swap with strike $\tilde{R}_Y$. The random variable ν_T^R is the realized variance up to maturity.

Let (the realized variance up to time t) $(\tilde{\nu}_t^R)_{0 \leq t \leq T}$ be defined by

$$\tilde{\nu}_t^R := \begin{cases} 0 & \text{if } t = 0 \\ \frac{1}{t} \int\limits_0^t \tilde{\nu}_u \, du & \text{if } 0 < t \leq T. \end{cases} \tag{3.3.4}$$

Proposition 3.3.3. *The (discounted) price of Y at time t, $t \in [0, T]$, is*

$$\tilde{Y}_t = \frac{t}{T} \tilde{\nu}_t^R + \frac{T - t}{T} \tilde{\nu}_t - \tilde{R}_Y. \tag{3.3.5}$$

Furthermore,

$$\tilde{R}_Y = \nu_0.$$

Proof. Let $t \in (0,T]$ then

$$\begin{aligned}
E^{\star}\left[\nu_T^R|\mathcal{F}_t\right] &= \frac{1}{T}\int_0^T E^{\star}\left[\tilde{\nu}_u|\mathcal{F}_t\right]du \\
&= \frac{1}{T}\left(\int_0^t E^{\star}\left[\tilde{\nu}_u|\mathcal{F}_t\right]du + \int_t^T E^{\star}\left[\tilde{\nu}_u|\mathcal{F}_t\right]du\right) \\
&= \frac{1}{T}\left(t\frac{1}{t}\int_0^t \tilde{\nu}_u du + \int_t^T \tilde{\nu}_t du\right) \\
&= \frac{1}{T}\left(t\tilde{\nu}_t^R + (T-t)\tilde{\nu}_t\right) \\
&= \frac{t}{T}\tilde{\nu}_t^R + \frac{T-t}{T}\tilde{\nu}_t.
\end{aligned}$$

For $t=0$ it follows that

$$\tilde{R}_Y = E^{\star}\left[\tilde{\nu}_T^R|\mathcal{F}_0\right] = \nu_0.$$

□

Corollary 3.3.4. *Let $Y_t := \tilde{Y}_t B_t$, $t \in [0,T]$, be the price of Y then*

$$Y_t = \frac{t}{T}\nu_t^R + \frac{T-t}{T}\nu_t - \frac{B_t}{B_T}R_Y, \quad t \in [0,T] \tag{3.3.6}$$

with $R_Y := B_T\tilde{R}_Y$ and

$$\nu_t^R := \frac{B_t}{t}\int_0^t \frac{\nu_u}{B_u}du, \quad t \in (0,T],$$

$\nu_0^R = 0$.

3.4. Problem discretization by the hedged Monte Carlo method

The discretization of the problem formulated in §3.2 by the hedged Monte Carlo method amounts to choose the process $F = (F_k)_{k=0}^K$ which generates the filtration $(\mathcal{F}_k)_{k=0}^K$, the basis size $(D_k)_{k=0}^{K-1}$ and the basis functions $(\{b_{k,d}\}_{d=1}^{D_k})_{k=0}^{K-1}$, associated functions $(f_k)_{k=0}^{K-1}$ and the number of Monte Carlo drawings I.
We set the stage for the calibration of $(d_k, \{b_{k,d}\}_{d=1}^{D_k})_{k=0}^{K-1}$ (see §3.4.3) in §3.4.1 and in §3.4.2.

3.4.1. The calibration framework

We have set $M = 2$ and $F = (F^1, F^2) = (S, \nu)$. Next we define the functions $f_k : F_{0\dots k}(\Omega) \subset \mathbb{R}^{(k+1)M} \to U_k$ and subsequently the class of basis functions $b_{k,j} : U_k \to \mathbb{R}$.

a) Since there are two sources of uncertainty, U_k should have at least two dimensions. The model is Markov and hence we set

$$f_k\left((S_0, \nu_0), \dots, (S_k, \nu_k)\right) \;:=\; (S_k, \nu_k), \quad k = 0, \dots, K-1$$

and

$$U_k = \begin{cases} \{(S_0, \nu_0)\} & \text{if } k = 0 \\ (0, \infty)^2 & \text{if } 1 \le k \le K-1. \end{cases}$$

Actually,

$$\begin{aligned} &P(\tilde{\Omega}_{1,k} := \{S_k \notin (0,\infty)\}) > 0 \quad \text{and} \\ &P(\tilde{\Omega}_{2,k} := \{\nu_k \notin (0,\infty)\}) > 0. \end{aligned}$$

If $\tilde{\Omega}_{\ell,k}$, $\ell = 1, 2$, has occurred, we discarded the realization and generated another one. We have never observed $\tilde{\Omega}_{1,k}$. We have observed $\bigcup\limits_{k=1}^{K} \tilde{\Omega}_{2,k}$ about once out of 10^4 realizations of ν.

b) We set $D_0 = 1$ and $D_k = D$ for $k = 1, \dots, K$. If $k = 0$, $b_{k,1} = 1$ and otherwise

$$b_{k,d} = \begin{cases} L^w_{d_1,d_2} & \begin{cases} d_1 = 0, \dots, D_1 - 1 & \text{if } D_2 = 0 \\ d_2 = 1, \dots, D_2 - 1 & \text{if } D_1 = 0 \\ 1 \le d_1 \le d_2,\ d_1 + d_2 \le \min\{D_1, D_2\} & \end{cases} \\ \text{or} & \\ N_{d_1,d_2,\ell_1,\ell_2} & \begin{cases} d_j = 1, \dots, D_j,\ j = 1, 2 \end{cases} \end{cases}$$

where $L^w_{d_1,d_2}$ is the (d_1, d_2)-th weighted Laguerre polynomial (see appendix A.1.1 for definition) and $N_{d_1,d_2,\ell_1,\ell_2}$, $\ell_j \in \{2, 3\}$, $j = 1, 2$, is the (d_1, d_2)-th B-spline function of order (ℓ_1, ℓ_2) (see appendix A.1.2 for definition). The parameters $D_j \in \mathbb{N}$, $j = 1, 2$, have to be chosen and determine D.

The splines $N_{d_1,d_2,\ell_1,\ell_2}$ are defined with respect to a grid

$$\left\{(x_{k,1,j_1}, x_{k,2,j_2}) \,|\, x_{k,1,j_1}, x_{k,2,j_2} \in \mathbb{R},\ j_1 = 1, \dots, n_1,\ j_2 = 1, \dots, n_2\right\},$$

where $n_1, n_2 \in \mathbb{N}$. The nodes x_{k,ℓ,j_ℓ}, $j_\ell = 1, \dots, n_\ell$, $\ell = 1, 2$, are chosen in the same manner as explained in §2.3.2; $x_{k,1,\cdot}$ with respect to S_k and $x_{k,2,\cdot}$ with respect to ν_k.

3.4.2. Implementation of the hedged Monte Carlo algorithm

The least squares problem (1.4.1) is solved by QR-factorization by means of Householder transformations.

All numerical results presented in this chapter have been obtained on a Intel(R) Core(TM)2 Duo CPU E8200 @ 2.66GHz processor using a (single) core with cpu MHz 1998.000 and cache size 6144 KB. The code has been written in `C++`.

3.4.3. Calibration of discretization parameters

The objective of this section is to determine an appropriate basis $\{b_{k,d}\}_{d=0}^{D}$. The computing time of the hedged Monte Carlo algorithm depends strongly on the size of D. We have observed that approximately

$$\tau \propto D^{\alpha}$$

for some α, $1.5 < \alpha < 2$. Hence it is very important to choose D as small as possible. Moreover, the choice of $\{b_{k,d}\}_{d=0}^{D}$ has strong influence on the approximation and on the consistency error on the one side and on the stochastic error on the other side. We determine an appropriate basis by the following three criteria:

$$E[\mathcal{R}^{\mathtt{hmc}}] \to \min, \tag{3.4.1a}$$

$$\mathrm{Var}[\mathcal{R}^{\mathtt{hmc}}] \to \min \tag{3.4.1b}$$

and

$$\tau^{\mathtt{hmc}} \to \min\,. \tag{3.4.1c}$$

Remark 3.4.1. The criteria are slightly different to (2.3.1) since we have no appropriate substitute for the locally risk-minimizing strategy $(V^{\mathtt{lrm}}, \vartheta^{\mathtt{lrm}})$. It would have been possible to consider the continuous time counterpart of the model here. It is complete and we could have computed the price $(H_k)_{k=0}^{K}$ of the hedging objective then by solving for instance a partial differential equation. Then we could have replaced $V_k^{\mathtt{lrm}}$ by H_k and $\vartheta_k^{1\,\mathtt{lrm}}$ by $\frac{\partial}{\partial S} H_k$. However it is not clear how to substitute $\vartheta_k^{2\,\mathtt{lrm}}$. The variance swaps allow to hedge in part price risk and in part volatility risk but we do not know to which extend these risks are hedged.

3.4.3.1. Numerical results

Table 3.4.1 shows the results obtained using various choices of the basis

$$\{b_{k,1}, \ldots, b_{k,D}\}, \quad k = 0, \ldots, K-1.$$

of A_k. The results obtained using Laguerre polynomials can be found in subtable (a). The results obtained using B-splines (surface splines) of order (2,2), (2,3), (3,2) and (3,3) can be found in subtables (b)-(e). A very good basis seems to be $(10, 2)$ surface splines of order $(3, 2)$. The results obtained there have been used as reference for the other results.

The results obtained using Laguerre polynomials seem to be of inferior quality than the reference results. The stochastic error became particular large (see column $\mathrm{s}_D^{\mathtt{hmc}}/\mathrm{s}^r$ for $\mathcal{R}$) when for the basis (D_1, D_2, D), D_1 was increased. The stochastic error seems to be the smallest for the basis $(3, 0, 3)$. In this case however the residual risk is almost five time larger than in the reference case. The results obtained using surface splines are much better than those from Laguerre polynomials. The ratio $\mathrm{m}^{\mathtt{hmc}}/\mathrm{m}^r$ for $\mathcal{R}$ has always been less than 1.51 while using the $(3, 0, 3)$ Laguerre basis the ratio has been 4.80.

Note that the Laguerre bases $(D_{x_1}, 0, D)$ and the spline bases $(D_{x_1}, 0)$ consist of no basis functions which depend on ν. This corresponds to the case

$$f_k((S_0, \nu_0), \dots, (S_k, \nu_k)) = (S_k), \quad k = 0, \dots, K-1,$$

i.e. to the case any ν-dependence is ignored. The portfolio value V_k is only a function of the price of the underlying asset, i.e. $V_k = V_k(s)$; analog the first and the second hedge, $\vartheta_k^1 = \vartheta_k^1(s)$ and $\vartheta_k^2 = \vartheta_k^2(s)$. If $H_k = H_k(s, \nu)$ is the market price of the hedging objective, then the value of the two portfolios together is

$$-H_k(s, \nu) + V_k(s), \quad s, \nu > 0.$$

This shows that the sensitivity of $-H_k(s, \nu)$ with respect to ν cannot be reduced with the hedge portfolio V_k. Both hedging instruments are used to hedge price risk in this case.

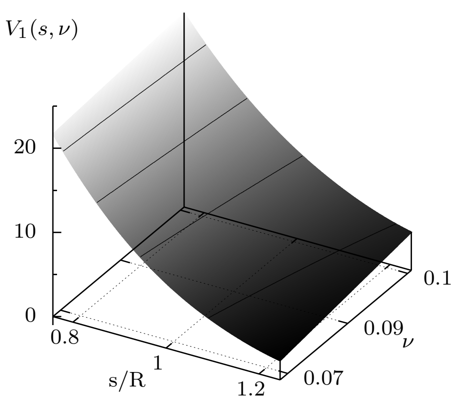
The results showed that the benefit of hedging in part price and in part volatility risk is a reduction of $\mathrm{m}^{\mathtt{hmc}}/\mathrm{m}^r$ for $\mathcal{R}$ from 1.39 (see (10,0) surface splines of order (3,2)) to 1 (see (10,2) surface splines of order (3,2)).

3.4.3.2. Visualization of a hedging strategy

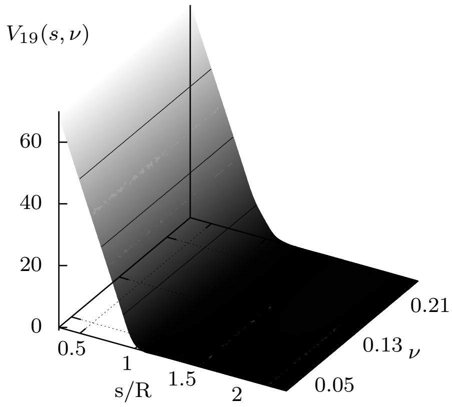
Figure 3.4.1 shows the regressed portfolio value V_k and the first hedge ϑ_{k+1}^1 and the second hedge ϑ_{k+1}^2 for $k = 1, 19$. In Figure 3.4.2 we have plotted the realizations of (S, ν) used at time $k = 1$ and $k = 19$. This shows where the density of (S_k, ν_k) is high and hence where the regression is better respectively worse (due to least squares). The portfolio value and the first hedge seem very well regressed. The second hedge seems to be most difficult to regress (like in the case of one source of uncertainty, cf. Figure 2.3.1 (e) and (f)).

3.5. Hedging under local risk minimization

The hedged Monte Carlo algorithm allows to compute a trading strategy which is an approximation to a locally risk-minimizing strategy. The local risk is defined as the squared cost increment in the framework of local risk minimization. The residual risk is not zero and hence the associated cost distribution is non-singular. Intuitively, it is clear that the distribution is the more peaked the more often one can hedge and/or the more hedging instruments are used. In which case, however,



(a) Level lines are at 5, 10, 15 and 20.



(b) Level lines are at 10, 30 and 50

Figure 3.4.1.: The plots show a hedging strategy obtained by the hedged Monte Carlo algorithm. We have set $I = 5000$ and A_k, $k > 0$, has been spanned by $(10, 2)$ B-splines of order $(3, 2)$; there have been $K = 20$ hedging opportunities; the S- and ν-range of the plots correspond to the domain of the B-splines;

does the cost distribution have a smaller second moment: if one hedges 10 times with two instruments or if one hedges 100 times with a single instrument?

In this section we analyze the local and global (residual) risk (see §3.5.2 and §3.5.1) and the cost distribution (see §3.5.3) if the hedging strategy is determined by the hedged Monte Carlo algorithm. The analysis is comparative and based on simulation.

Given $(V^{\mathtt{hmc}}, \vartheta^{\mathtt{hmc}})$ we generate another set of realizations $F^{(I+1)}, \dots, F^{(2I)}$ of F and compute the local, $\Delta C^{\mathtt{hmc}}_{k+1}$, and the global, $\Delta C^{\mathtt{hmc}}$, cost increment. This gives rise to a realization of the local $\mathcal{R}_k$ and the global $\mathcal{R}$ risk estimator, cf. §1.5.

We compare the results for various choices of I (the number of simulated paths) and K (the number of hedging opportunities).

The test problem is basically that one formulated in §3.2. We either set

(a) $X = (X^1)$, i.e. there is only one hedging instrument, the underlying asset

or we set

(b) $X = (X^1, X^2)$, i.e. the hedging instruments are the underlying asset and the variance swaps.

We will refer to (a) as simple hedging or the simple hedged case and to (b) as double hedging or the double hedged case.

3.5.1. Global risk

Table 3.5.1 documents the mean and the standard deviation of the $\mathcal{R}$-estimator for I ($2500 \uparrow 20000$) paths and $K = 20$ hedging opportunities (subtable (a)) and for K ($10 \uparrow 100$) and $I = 5000$ (subtable (b)).

Subtable (a) shows that increasing the number of paths (I) by factor 8 (from 2500 to 20000) leads to a reduction of the standard deviation of $\mathcal{R}$ by factor $3.10 = 1/0.3222$ in the simple hedged case and $2.69 = 1/0.3715$ in the double hedged case. Compared to $\sqrt{8} = 2.83$ we see that the reduction rate is about $1/2$ which is typical for Monte Carlo methods [Fis96, Jae02, Gla03].
Increasing I leads as well to slightly less risky hedging strategies ($\mathrm{m}(\mathcal{R})$ decreases, see subtable (a)). This shows that the approximation of the `hmc` algorithm is the better the larger I.

Subtable (b) shows that increasing the number of hedging opportunities (K) has a much stronger risk reduction effect. Increasing K by factor 10 (from 10 to 100) leads to a reduction of the mean of $\mathcal{R}$ by factor $2.36 = 1/0.4237$ in the simple hedged case and $3.21 = 1/0.3115$ in the double hedged case. The reduction rate is thus about $1/2$ in the double hedged case and less than $1/2$ in the simple hedged case ($\sqrt{10} = 3.16$).
In view of the standard deviation of $\mathcal{R}$ we see that it decreases by factor $2.60 = 1/0.3839$ (simple hedging) and by $4.40 = 1/0.2275$ (double hedging) for K from 10 to 100. The reason that the standard deviation of $\mathcal{R}$ decreases at all is that risk minimization leads at the same time to variance reduction.

Instead of hedging more frequently, hedging with more instruments seems to be more useful for risk reduction. In the double hedged case, the residual risk is about one fourth of the residual risk in the simple hedged case (for an equal number of hedging opportunities), see the last column in subtables (a) and (b). One can see as well that 100 times simple hedging leaves more risk ($\mathrm{m}(\mathcal{R}) = 3.1345 \times 0.4237 = 1.3280$) than 10 times double hedging ($\mathrm{m}(\mathcal{R}) = 0.9525$).

3.5.2. Local risk

Figure 3.5.1 shows on left (subfigure (a)) the local risk $\mathcal{R}_k$ for $k = 0, \ldots, K-1$. One can see that in the double hedged case $\mathcal{R}_k$ is much smaller than in the simple hedged case for $k = 0, \ldots, 3/4K$. However, $\mathcal{R}_k$ increases much stronger in the double hedged case close to maturity than in the simple hedged case. This effect is visualized on the right (subfigure (b)).

In the simple hedged case, $X = (X^1)$, the increments

$$\begin{aligned} \Delta C^{\mathbf{hmc}}_{k+1} = & V^{\mathbf{hmc}}_{k+1}(f_{k+1}(F_{0\ldots k+1})) - V^{\mathbf{hmc}}_{k}(f_k(0\ldots F_k)) \\ & - \vartheta^{\mathbf{hmc}}_{k+1}(f_k(F_{0\ldots k}))\left(X_{k+1}(F_{0\ldots k+1}) - X_k(F_{0\ldots k})\right), \\ & \qquad\qquad k = 0, \ldots, K-1 \end{aligned} \tag{3.5.1}$$

are pairwise independent and hence we expect that

$$\mathcal{R}^2 \approx \sum_{k=0}^{K-1} \mathcal{R}_k^2,$$

cf. §1.5. If $\mathcal{R}^2 = \sum_{k=0}^{K-1} \mathcal{R}_k^2$, it is optimal to have equal residual risk on each hedging period, i.e. $\mathcal{R}_k^2 = \mathcal{R}^2/K$. This would correspond to a straight line with slope 1 in subfigure (b).

In the double hedged case, $X = (X^1, X^2)$, X^2 does not have independent increments, cf. (3.2.3b), and hence the increments $\Delta C_{k+1}^{\mathtt{hmc}}$ are independent neither.

3.5.3. The distribution of ΔC - higher moments

Yet, our analysis has focused on the second moment of the cost (increments), ΔC and ΔC_k, as this is the way we have defined $\mathcal{R}$ and $\mathcal{R}_k$. Higher moments however exist and should receive some attention.

Figure 3.5.2 shows on the left (subfigure (a)) the ΔC distribution from simple hedging and on the right (subfigure (b)) the ΔC distribution from double hedging. Table 3.5.2 reports the sample mean and the sample standard deviation of the mean, the standard deviation, the skewness and the kurtosis estimator of ΔC.
The results support what can be seen in Figure 3.5.2. The mean of ΔC is approximately zero for simple and for double hedging. Recall that locally risk-minimizing strategies are mean self-financing, cf. Theorem 1.2.7, and hence $E[\Delta C] = 0$. The distribution of ΔC has much smaller standard deviation in the case of double hedging than in the case of simple hedging. Compared to the normal distribution, the ΔC-distribution has high kurtosis in the case of double hedging. Note, however, that the standard deviation of the kurtosis estimator is rather high.

Knowledge of the full distribution of ΔC can be used for instance to compute the probability of losing a certain amount of money, cf. Figure 3.5.3.
The plot is quite impressive: imagine you have sold H for V_0 then you will lose with probability 30.0% more than 10% of V_0 in the case of simple hedging. In the double hedged case, however, the probability is only 2.7%.

3.6. Efficiency and speed-up of the hedged Monte Carlo algorithm

We start with introducing a measure of efficiency.

Let θ be a Monte Carlo estimator, i.e. a random variable and realizations of θ are obtained by Monte Carlo simulation. A standard measure for the efficiency of a Monte Carlo method is

$$\varepsilon := s(\theta)^2 \times \tau, \tag{3.6.1}$$

where $s(\theta)$ is the sample standard deviation of θ and τ is the computing time for one realization of θ. The definition can be motivated by the central limit theorem. We refer to [BBG97, Gla03].

In order to measure the efficiency of the hedged Monte Carlo algorithm we set $\theta = V_0$.
The initial portfolio value is a natural candidate for the price of the hedging objective. The continuous counterpart of the model introduced in section 3.2 is complete and hence in this framework the hedging objective has a unique price. The price can be approximated by standard Monte Carlo simulation under the risk neutral measure. For this reason we compare the efficiency of the hedged Monte Carlo algorithm with the efficiency of standard risk neutral Monte Carlo. We will denote the Monte Carlo estimator of the price in the risk neutral framework as well by V_0.

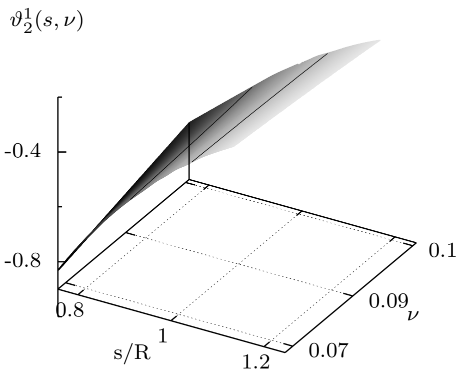
The speed-up of the hedged Monte Carlo algorithm with respect to risk neutral Monte Carlo is

$$\varsigma := \frac{\varepsilon^{\mathtt{hmc}}}{\varepsilon^{\mathtt{rnmc}}}. \tag{3.6.2}$$

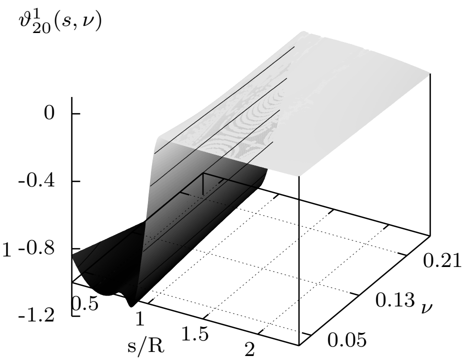
Note that for equal standard deviation the computing time of the risk neutral Monte Carlo algorithm is ς-times the computing time of the hedged Monte Carlo algorithm. Since the computing time of a Monte Carlo algorithm depends linearly on the number of simulated paths, one has to simulate ς-times as many paths using risk neutral Monte Carlo than using the hedged Monte Carlo algorithm.

Note further that the risk neutral Monte Carlo algorithm and hedged Monte Carlo algorithm are two different algorithms with two different areas of application in general. The first one for pricing and the second one for hedging. The initial portfolio value is not equal to the price of the hedging objective in the risk neutral framework. The reason is that the risk neutral measure and the minimal martingal measure are different. Speed-up is only what it has been defined.

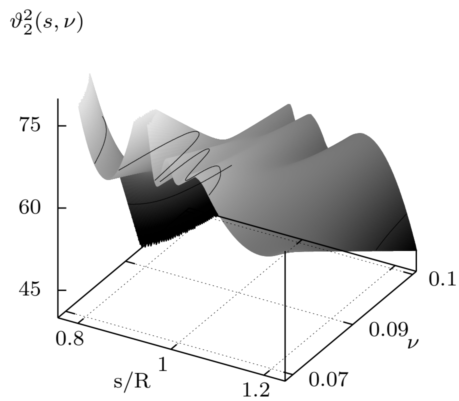
Like in the preceding section we have generated numerical results for simple hedging ($X = (X^1)$) and for double hedging ($X = (X^1, X^2)$). The performance of the hedged Monte Carlo algorithm has been investigated for I $(2500 \uparrow 20000)$. The numerical results are presented in Table 3.6.1. One can see that the standard deviation of the risk neutral Monte Carlo V_0-estimator is much larger than the standard deviation of the hedged Monte Carlo V_0-estimator. For the computing time it is vice versa. The hedged Monte Carlo algorithm is about as efficient as the risk neutral Monte Carlo if $X = (X^1)$, but highly efficient if $X = (X^1, X^2)$. In the last case the speed-up factor is about seven.



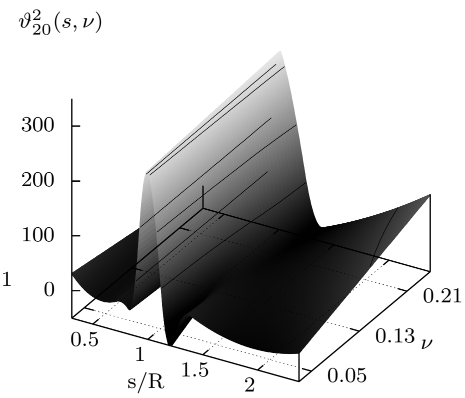
(c) Level lines are at -0.5 and -0.3.



(d) Level lines are at -0.9, -0.7, ...



(e) Level lines are at 50, 70 and 90.



(f) Level lines are at 50, 150 and 250.

Figure 3.4.1.: The plots show a hedging strategy obtained by the hedged Monte Carlo algorithm. We have set $I = 5000$ and A_k, $k > 0$, has been spanned by $(10, 2)$ B-splines of order $(3, 2)$; there have been $K = 20$ hedging opportunities; the S- and ν-range of the plots correspond to the domain of the B-splines;

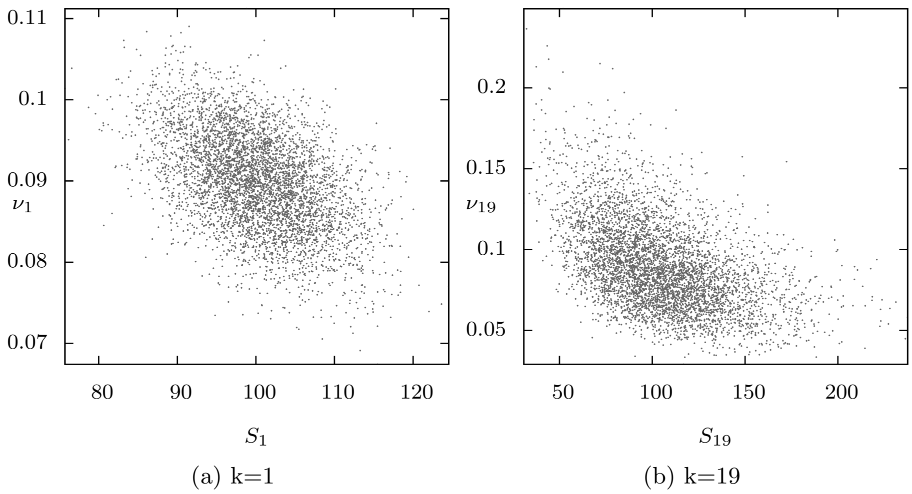


Figure 3.4.2.: The 5 000 realisations of (S_k, ν_k) used by the hedged Monte Carlo algorithm to compute the hedging strategy which is partially visualized in Figure 3.4.1.

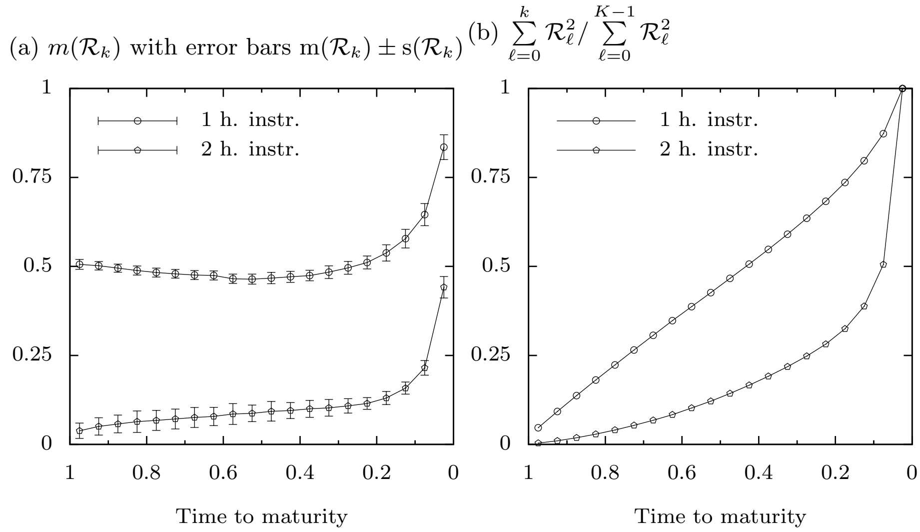


Figure 3.5.1.: Local risk versus time to maturity. '1 h. instr.' refers to the case $X = (X^1)$ and '2 h. instr.' refers to the case $X = (X^1, X^2)$; X^1 is defined in (3.2.3) and X^2 is defined in (3.2.3b). There has been 20 hedging opportunities ($K = 20$). The data has been obtained by $I = 5000$ and A_k, $k > 0$, spanned by $(10, 2)$ b-splines of order $(3, 2)$. The sample estimators m and s are based on 250 samples.

(a) Laguerre polynomials

d	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	
(3,0,3)	-0.0153	4.2660	4.7960	2.5714	0.11
(3,2,4)	-0.0213	4.1702	4.7356	2.6000	0.14
(3,3,6)	-0.0183	3.9362	4.4489	5.0612	0.18
(3,4,7)	-0.0174	3.8723	4.4292	7.2612	0.21
(3,5,8)	-0.0171	3.8404	4.4152	7.3918	0.24
(4,0,4)	-0.0139	3.9787	4.5481	5.9837	0.13
(4,2,5)	-0.0196	3.8936	4.5078	6.0531	0.16
(4,3,7)	-0.0183	3.8936	4.4405	5.4327	0.21
(4,4,10)	-0.0017	4.2128	4.3515	51.4408	0.31
(4,5,11)	-0.0014	4.2234	4.3923	53.1143	0.36
(5,0,5)	0.0028	3.9681	4.2766	40.7551	0.15
(5,2,6)	-0.0030	3.9362	4.2240	41.0204	0.18
(5,3,8)	-0.0028	3.9255	4.2119	41.0449	0.24
(5,4,11)	0.0000	3.4043	3.7955	38.7061	0.36
(5,5,15)	-0.0041	3.4468	3.6296	49.9388	0.56
(6,0,6)	0.0008	3.1170	3.3983	37.8163	0.17
(6,2,7)	-0.0052	3.0426	3.3308	38.1918	0.21
(6,3,9)	-0.0051	3.0426	3.3264	38.4980	0.28
(6,4,12)	-0.0034	3.0106	3.3116	38.0408	0.40
(6,5,16)	0.0094	16.2660	6.3000	383.3429	0.62
(7,0,7)	0.0146	17.2234	7.2050	446.8408	0.20
(7,2,8)	0.0086	17.2660	7.1573	448.9633	0.24
(7,3,10)	0.0086	17.2553	7.1464	448.4408	0.32
(7,4,13)	0.0104	17.2234	7.1696	448.0490	0.45
(7,5,17)	0.0100	17.6277	7.3167	460.4612	0.69

Table 3.4.1.: Sample mean (m) and sample standard deviation (s) of the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator using 250 samples (column 2-5); the superscripts ($\texttt{hmc}$,r) refer to the underlying hedging strategy of the $\mathcal{V}_0$- and $\mathcal{R}$-estimator; r refers to the reference strategy; $\tau_D^{\texttt{hmc}}$ denotes the computing time of the hedged Monte Carlo algorithm for (V, ϑ); We have set $I = 5000$, $K = 20$ and A_k is spanned by different bases: the heading of the subtable indicate the basis type; the basis size and composition is specified in the first column: for Laguerre polynomials the triple is (D_1, D_2, D) and for B-splines the pair is (D_1, D_2) which gives $D = D_1 \times D_2$.

(b) b-splines of order (2,2)

d	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	
(8,0)	0.0072	1.3511	1.5078	0.8612	0.26
(8,2)	0.0026	1.0957	1.1374	0.9143	0.67
(8,3)	0.0028	1.0957	1.1457	0.9592	1.31
(8,4)	0.0031	1.0851	1.1600	1.0000	2.20
(10,0)	0.0061	1.3085	1.4420	0.9184	0.33
(10,2)	0.0014	1.0213	1.0588	1.0449	0.96
(10,3)	0.0019	1.0000	1.0699	1.2204	1.95
(10,4)	0.0021	1.0106	1.0866	1.2490	3.38
(15,0)	0.0056	1.2872	1.4100	0.9184	0.59
(15,2)	0.0011	1.0000	1.0190	1.1020	1.97
(15,3)	0.0014	0.9894	1.0367	1.2122	4.28

(c) b-splines of order (2,3)

d	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\texttt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\texttt{hmc}} - \mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	$\mathrm{m}_D^{\texttt{hmc}}/\mathrm{m}^r$	$\mathrm{s}_D^{\texttt{hmc}}/\mathrm{s}^r$	
(8,0)	0.0072	1.3511	1.5078	0.8612	0.26
(8,3)	0.0029	1.0851	1.1545	1.0000	1.32
(8,4)	0.0033	1.0851	1.1804	1.2163	2.21
(10,0)	0.0061	1.3085	1.4420	0.9184	0.33
(10,3)	0.0020	1.0106	1.0840	1.6204	1.96
(10,4)	0.0023	1.0319	1.1142	1.9714	3.41
(15,0)	0.0056	1.2872	1.4100	0.9184	0.59
(15,3)	0.0017	1.0106	1.0615	1.8449	4.28
(15,4)	0.0020	1.0851	1.1088	3.6898	7.93

Table 3.4.1.: Sample mean (m) and sample standard deviation (s) of the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator using 250 samples (column 2-5); the superscripts ($\texttt{hmc}$,r) refer to the underlying hedging strategy of the $\mathcal{V}_0$- and $\mathcal{R}$-estimator; r refers to the reference strategy; $\tau_D^{\texttt{hmc}}$ denotes the computing time of the hedged Monte Carlo algorithm for (V, ϑ); We have set $I = 5000$, $K = 20$ and A_k is spanned by different bases: the heading of the subtable indicate the basis type; the basis size and composition is specified in the first column: for Laguerre polynomials the triple is (D_1, D_2, D) and for B-splines the pair is (D_1, D_2) which gives $D = D_1 \times D_2$.

(d) b-splines of order (3,2)

d	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\mathtt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\mathtt{hmc}} - \mathrm{m}^r$	$\mathrm{s}_D^{\mathtt{hmc}}/\mathrm{s}^r$	$\mathrm{m}_D^{\mathtt{hmc}}/\mathrm{m}^r$	$\mathrm{s}_D^{\mathtt{hmc}}/\mathrm{s}^r$	
(8,0)	0.0041	1.3191	1.4443	0.7429	0.29
(8,2)	-0.0003	1.0426	1.0766	0.8735	0.70
(8,3)	0.0000	1.0638	1.0899	1.2857	1.34
(8,4)	0.0002	1.0957	1.1230	2.2286	2.22
(10,0)	0.0045	1.2872	1.3904	0.8612	0.37
(10,2)	0.0000	1.0000	1.0000	1.0000	1.00
(10,3)	0.0000	0.9894	1.0136	1.1592	1.98
(10,4)	0.0000	1.0426	1.0426	1.6816	3.42
(15,0)	0.0048	1.2553	1.3884	0.9143	0.64
(15,2)	0.0004	0.9574	0.9944	1.0980	2.01
(15,3)	0.0005	0.9574	1.0157	1.5551	4.33
(15,4)	0.0005	1.0638	1.0933	6.0000	7.99

(e) b-splines of order (3,3)

d	$\mathcal{V}_0$		$\mathcal{R}$		$\tau_D^{\mathtt{hmc}}/\tau^r$
	$\mathrm{m}_D^{\mathtt{hmc}} - \mathrm{m}^r$	$\mathrm{s}_D^{\mathtt{hmc}}/\mathrm{s}^r$	$\mathrm{m}_D^{\mathtt{hmc}}/\mathrm{m}^r$	$\mathrm{s}_D^{\mathtt{hmc}}/\mathrm{s}^r$	
(8,0)	0.0041	1.3191	1.4443	0.7429	0.29
(8,3)	-0.0003	1.0638	1.0984	1.2653	1.34
(8,4)	-0.0002	1.1064	1.1430	2.6694	2.23
(10,0)	0.0045	1.2872	1.3904	0.8612	0.37
(10,3)	0.0000	1.0106	1.0241	1.3143	1.99
(10,4)	-0.0001	1.0638	1.0686	2.1714	3.42
(15,0)	0.0048	1.2553	1.3884	0.9143	0.64
(15,3)	0.0005	0.9681	1.0332	1.7837	4.33
(15,4)	0.0007	1.0426	1.0876	3.3143	7.99

Table 3.4.1.: Sample mean (m) and sample standard deviation (s) of the $\mathcal{V}_0$- and the $\mathcal{R}$-estimator using 250 samples (column 2-5); the superscripts ($\mathtt{hmc}$,r) refer to the underlying hedging strategy of the $\mathcal{V}_0$- and $\mathcal{R}$-estimator; r refers to the reference strategy; $\tau_D^{\mathtt{hmc}}$ denotes the computing time of the hedged Monte Carlo algorithm for (V, ϑ); We have set $I = 5000$, $K = 20$ and A_k is spanned by different bases: the heading of the subtable indicate the basis type; the basis size and composition is specified in the first column: for Laguerre polynomials the triple is (D_1, D_2, D) and for B-splines the pair is (D_1, D_2) which gives $D = D_1 \times D_2$.

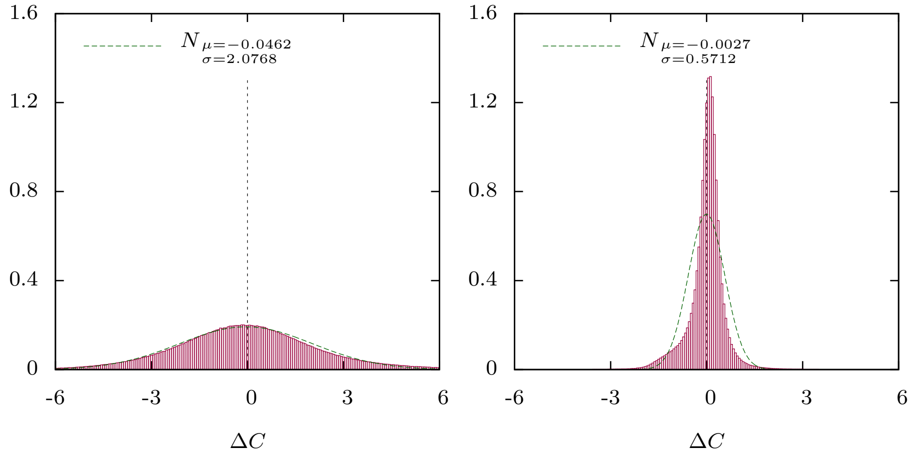
(a) Number of paths generated (column 1). For $I = 2500$ we have obtained for $\mathcal{R}$: $\mathrm{m}_I^{1h} = 2.3380$, $\mathrm{s}_I^{1h} = 0.0509$ and $\mathrm{m}_I^{2h} = 0.6104$, $\mathrm{s}_I^{2h} = 0.0358$

$I/10^3$	1 h. instr.: $\mathcal{R}$		2 h. instr.: $\mathcal{R}$		$\mathrm{m}_I^{2h}/\mathrm{m}_I^{1h}$
	$\mathrm{m}_{I=2500}/\mathrm{m}_I$	$\mathrm{s}_{I=2500}/\mathrm{s}_I$	$\mathrm{m}_{I=2500}/\mathrm{m}_I$	$\mathrm{s}_{I=2500}/\mathrm{s}_I$	
2.5	1.0000	1.0000	1.0000	1.0000	0.2611
5.0	0.9943	0.5914	0.9726	0.6844	0.2554
7.5	0.9909	0.5069	0.9615	0.5587	0.2533
10.0	0.9892	0.4499	0.9567	0.4665	0.2525
12.5	0.9883	0.4028	0.9522	0.3966	0.2515
15.0	0.9877	0.3870	0.9497	0.3603	0.2510
17.5	0.9871	0.3458	0.9472	0.3464	0.2505
20.0	0.9868	0.3222	0.9464	0.3715	0.2504

(b) Number of hedging opportunities (column 1). For $K = 10$ we have obtained for $\mathcal{R}$: $\mathrm{m}_I^{1h} = 3.1345$, $\mathrm{s}_I^{1h} = 0.0409$ and $\mathrm{m}_I^{2h} = 0.9525$, $\mathrm{s}_I^{2h} = 0.0444$

K	1 h. instr.: $\mathcal{R}$		2 h. instr.: $\mathcal{R}$		$\mathrm{m}_I^{2h}/\mathrm{m}_I^{1h}$
	$\mathrm{m}_{I=2500}/\mathrm{m}_I$	$\mathrm{s}_{I=2500}/\mathrm{s}_I$	$\mathrm{m}_{I=2500}/\mathrm{m}_I$	$\mathrm{s}_{I=2500}/\mathrm{s}_I$	
10	1.0000	1.0000	1.0000	1.0000	0.3039
15	0.8350	0.8973	0.7545	0.7432	0.2746
20	0.7416	0.7359	0.6233	0.5518	0.2554
25	0.6768	0.6724	0.5424	0.4550	0.2435
30	0.6297	0.6039	0.4896	0.4167	0.2363
50	0.5231	0.5355	0.3783	0.2568	0.2197
75	0.4592	0.4279	0.3321	0.2230	0.2197
100	0.4237	0.3839	0.3115	0.2275	0.2234

Table 3.5.1.: Sample mean (m) and standard deviation (s) of the risk estimator, $\mathcal{R}$, for simple and double hedging, i.e. for $X = (X^1)$ (column 2 and 3) and for $X = (X^1, X^2)$ (column 4 and 5). Column 6 gives the ratio of the sample mean of $\mathcal{R}$ of 'hedging with two instruments' over 'hedging with one instrument'. The results have been obtained generating 250 samples. For each sample we have run the hedged Monte Carlo algorithm. We have used $(10, 2)$ B-splines of order $(3, 2)$ to approximate the hedging strategy.



(a) 1 hedging instrument

(b) 2 hedging instruments

Figure 3.5.2.: The plots show the estimated probability distribution of ΔC for simple hedging on the left (i.e. $X = (X^1)$) and for double hedging on the right (i.e. $X = (X^1, X^2)$). The plotted data are relative frequencies obtained by simulation. The locally risk-minimizing strategy has been approximated 250 times by the hedged Monte Carlo algorithm. For each sample I realizations of ΔC have been generated. The free discretization parameters I and the basis of A_k have been set to $I = 5000$ and A_k, $k > 0$, has been spanned by $(10, 2)$ B-splines of order $(3, 2)$. There has been 20 hedging opportunities ($K = 20$). If the option H is sold for V_0 and if a hedge portfolio is contructed according to (V, ϑ), one will make loss if $\Delta C > 0$ and profit if $\Delta C < 0$.

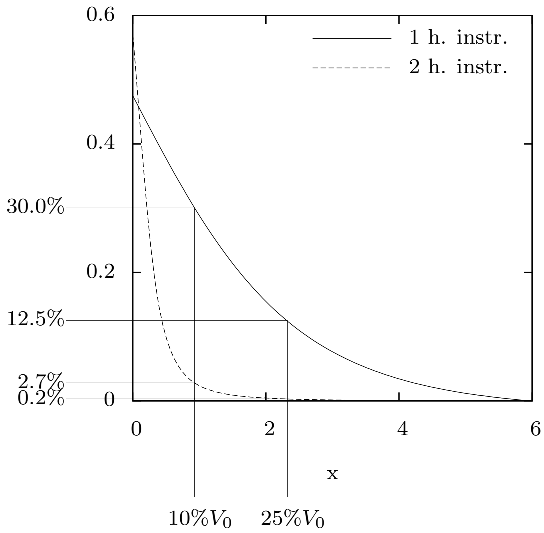


Figure 3.5.3.: Estimate of the cumulative distribution functions $x \mapsto P(\Delta C > x)$ computed from the estimated probability distribution of ΔC (plotted in figure 3.5.2).

Estimator	1 h. instr.		2 h. instr.	
	m	s	m	s
Mean(ΔC)	0.0363	0.0432	-0.0034	0.0121
StDev(ΔC)	2.3226	0.0304	0.5949	0.0259
Skew(ΔC)	0.4702	0.0821	-0.9422	1.5500
Kurt(ΔC)	1.6515	0.4206	29.8749	44.1677

Table 3.5.2.: Estimated mean, standard deviation, skewness and kurtosis of the distribution of ΔC for simple hedging on the left (i.e. $X = (X^1)$) and for double hedging on the right (i.e. $X = (X^1, X^2)$). The estimators (Mean, StDev, Skew and Kurt) are defined in appendix A.2. The sample mean (m) and the sample standard deviation (s) is based on 500 samples. A sample has been obtained running the hedged Monte Carlo algorithm and I-times simulation of ΔC. We have set $I = 5000$ and A_k, $k > 0$, has been spanned by $(10, 2)$ B-splines of order $(3, 2)$. There has been 20 hedging opportunities ($K = 20$).

I	rnmc		hmc1		hmc2	
	$s(V_0)$	τ/sec	$s(V_0)$	τ/sec	$s(V_0)$	τ/sec
2500	0.2893	0.03	0.0445	1.21	0.0120	2.35
5000	0.1951	0.07	0.0331	2.45	0.0094	4.72
7500	0.1582	0.10	0.0287	3.68	0.0072	7.11
10000	0.1312	0.13	0.0218	4.90	0.0057	9.58
12500	0.1202	0.17	0.0197	6.16	0.0052	12.67
15000	0.1097	0.21	0.0181	7.57	0.0046	16.25
17500	0.1028	0.25	0.0168	9.48	0.0043	20.95
20000	0.0987	0.30	0.0165	11.96	0.0042	26.69

(a) Sample standard deviation of the V_0 estimator (column 2, 4 and 6) and computing time τ for one realization (column 3, 5 and 7); the number of simulated paths I (column 1).

I	efficiency (ε)			speed up (ς)	
	rnmc	hmc1	hmc2	hmc1	hmc2
2500	0.002511	0.002396	0.000338	1.05	7.43
5000	0.002664	0.002684	0.000417	0.99	6.39
7500	0.002503	0.003031	0.000369	0.83	6.78
10000	0.002238	0.002329	0.000311	0.96	7.20
12500	0.002456	0.002391	0.000343	1.03	7.16
15000	0.002527	0.002480	0.000344	1.02	7.35
17500	0.002642	0.002676	0.000387	0.99	6.83
20000	0.002923	0.003256	0.000471	0.90	6.21

(b) Efficiency (column 2-4) and speed-up (column 5-6) for various I (number of simulated paths); ε is defined in (3.6.1) and ς is defined in (3.6.2).

Table 3.6.1.: rnmc refers to the risk neutral Monte Carlo algorithm and hmc refers to the hedged Monte Carlo algorithm; we have generated 500 samples for rnmc and 250 for hmc; hmc1 refers to the case $X = (X^1)$ and hmc2 refers to the case $X = (X^1, X^2)$. We have set $K = 20$ and we have used $(10, 2)$ B-splines of order $(3, 2)$.

Part II.

Optimal control of options

4. Optimal control of European double barrier basket options

This chapter is to a large extent identical to [HL11].

Abstract: We consider European double barrier basket call options on two underlyings with an upper and a lower knock-out barrier. The payoff function is stripewise affine linear. The exact shape of the payoff function and the level of the rebate are determined by parameters (controls) that have to be chosen such that the delta of the option is as close as possible to a predefined constant profit/loss. This leads to a control constrained optimal control problem for the two-dimensional Black-Scholes equation with Dirichlet boundary control and finite time control. Based on the variational formulation of the problem in an appropriate Sobolev space setting, we prove the existence of a unique solution and state the first order necessary optimality conditions. A semi-discretization in space by conforming P1 finite elements with respect to a simplicial triangulation of the computational domain gives rise to a semi-discrete control constrained optimal control problem for a linear system of first order ordinary differential equations. A further discretization in time by the backward Euler scheme results in a fully discrete optimization problem that is solved numerically by the projected BFGS method with Armijo line search. Numerical examples for some selected test cases illustrate the benefits of hedging with European double barrier basket options in case of optimally controlled cash settlements.

Contents

4.1. Introduction

Barrier options are exotic options and mostly traded over-the-counter. There are various kinds of barrier options on the market: barrier options with single or double barrier, with one or more underlying assets, barrier options which are activated/deactivated if a barrier is crossed, with and without rebate at the barrier etc.

The problem which is investigated in this chapter is to create a double barrier option on a basket of two assets such that it has (almost) constant delta.

Such an option would be appealing i) for someone who holds a portfolio of the underlying assets and wants to have protection against (decreasing) asset prices and ii) for someone who holds a portfolio of this option and wants to hedge it.

The valuation of a single barrier option on one underlying has already been dealt with in the seminal paper by Merton [Mer73] and subsequently studied in [Car95, CV96, RR91, Rit98]. The first contribution addressing barrier options with more than one underlying is [HK94]. In particular, the authors are concerned with barrier options on a single stock with the barrier being determined by another asset, whereas valuation formulas for barrier options on a basket have been derived later in [KWY98, WK03]. Hedging techniques for barrier options have been considered by different approaches including static hedging based on the equivalence to contingent claims with specifically adjusted payoff functions [CC97, CEG98, Pou06], the partial differential equation (PDE) formulation [AAE02, DEK95, LYH03, NP06, Pou06, Sbu05], and stochastic optimization [GM07, Mar09b, MS09].

We follow an optimal control approach for the optimization of European double barrier basket option. The paper is organized as follows: In section 4.2, we formulate the problem. The payoff function is parametrized by a finite number of parameters $u = (u_1, \ldots, u_M)^T$, $M \in \mathbb{N}$, featuring the control. The rebate at the upper barrier is u_M. The control variables are bilaterally constrained and have to be chosen in such a way that a tracking type objective functional in terms of the delta of the option is minimized. The state variable is the Black-Scholes price of the option.
This amounts to the solution of a control constrained optimal control problem for the Black-Scholes equation in some space-time domain $Q := \Omega \times (0, T), T > 0$, where Ω is a trapezoidal domain in $\mathbb{R}^2$ determined by the lower and upper barriers K_{min} and K_{max}. The rebate at the upper barrier enters as a Dirichlet boundary control, whereas the payoff enter as a final time control vector. A particular feature is that the Dirichlet boundary conditions on the boundaries parallel to the coordinate axes are given by the solution of associated 1D Black-Scholes equations.
In section 4.3, we reformulate the problem as an initial control/Dirichlet boundary control problem by means of a simple transformation in time and deal with its weak formulation in a weighted Sobolev space setting. Section 4.4 is devoted to the derivation of the first order necessary optimality conditions involving adjoint states that satisfy backward in time parabolic PDEs as well as a variational inequality due to the bilateral constraints on the control. In section 4.5, we are concerned with

a semi-discretization in space by conforming P1 finite elements with respect to a simplicial triangulation of the computational domain. The semi-discrete optimal control problem amounts to the minimization of a semi-discrete objective functional subject to systems of first order ordinary differential equations (ODEs) obtained by the finite element approximation in space and the constraints imposed on the controls. It thus represents a control constrained initial control problem for the respective systems of first order ODEs in terms of the associated mass and stiffness matrices as well as input matrices expressing the input from the semi-discretized boundary controls at the upper barrier. The corresponding semi-discrete optimality system nicely reflects the intrinsic couplings between the states, their adjoint counterparts, and the controls. Using a further discretization in time by the implicit Euler method, the resulting fully discrete optimality system is studied in section 4.6. Its numerical solution is realized by the projected BFGS method with Armijo line search. Finally, section 4.8 contains a documentation of numerical results illustrating the application of the optimal control approach.

Optimal control of double barrier options on a single underlying asset has been investigated before [BB02]. Their work however is quite different to ours: they do not consider a weak formulation of the Black-Scholes equation; they use finite differences for the discretization in space and the Crank-Nicholson method for the discretization in time, cf. [AP05, Hul08]. The resulting discrete optimal control problem is then solved by sequential quadratic programming, cf. [GMW82, NW99]. The gradient of the objective functional has been computed through automatic differentiation, cf. [AP05, Gri00].

4.2. The optimal control problem

We consider a European double barrier basket call option on a basket consisting of two assets with prices S_i, $1 \leq i \leq 2$. We assume that the option has maturity T, $T \in \mathbb{R}_+$, strike K, $K \in \mathbb{R}_+$, and barriers K_{min} and K_{max}, satisfying K_{min}, $K_{max} \in \mathbb{R}_+$, $K_{min} < K < K_{max}$.

We consider the Black-Scholes model. Hence the price of the option, y_t, at time t is unique and a function of t and the prices, $S_{i,t}$, $1 \leq i \leq 2$, of the underlying assets at time t, i.e. $y_t = y(S,t)$, $S = (S_1, S_2)$. The temporal domain for the price of the option y is $[0,T]$ and spatial domain is given by the trapezoid

$$\Omega := \{S = (S_1, S_2) \in \mathbb{R}_+^2 \mid K_{min} < |S| < K_{max}\}, \tag{4.2.1}$$

where $|S| := S_1 + S_2$, with boundaries

$$\begin{aligned} \Gamma_1 &:= (K_{min}, K_{max}) \times \{0\}, & \Gamma_2 &:= \{0\} \times (K_{min}, K_{max}), \\ \Gamma_3 &:= \{S \in \mathbb{R}_+^2 \mid |S| = K_{min}\}, & \Gamma_4 &:= \{S \in \mathbb{R}_+^2 \mid |S| = K_{max}\}. \end{aligned} \tag{4.2.2}$$

Let us denote by $r = r(t)$, $t \in [0,T]$, the risk-free interest rate and by $\sigma_k = \sigma_k(S,t)$, $1 \leq k \leq 2$, $S \in \Omega$, $t \in [0,T]$, the volatilities of the assets. Moreover, we refer to

$\rho = (\rho_{k\ell})_{k,\ell=1}^2$ with $\rho_{kk} = 1$, $1 \le k \le 2$, and $\rho_{12} = \rho_{21} = 2\rho/(1+\rho^2)$, $-1 < \rho < +1$, as the correlation matrix. We set $\xi = (\xi_{k\ell})_{k,\ell=1}^2$ where $\xi_{k\ell} := \rho_{k\ell}\sigma_k\sigma_\ell$, $1 \le k, \ell \le 2$. It is well-known (cf., e.g., [AP05, BB02, Top98]) that the price y_Q, $Q := \Omega \times (0,T)$, of the option satisfies the following boundary value problem for the Black-Scholes equation with a final time condition at maturity T:

$$\frac{\partial y_Q}{\partial t} + L_\Omega(t) y_Q = 0 \qquad \text{in } Q := \Omega \times (0,T), \tag{4.2.3a}$$

$$y_Q = y_{\Sigma_j} \qquad \text{on } \Sigma_j := \Gamma_j \times (0,T)\ ,\ 1 \le j \le 4, \tag{4.2.3b}$$

$$y_Q(\cdot, T) = y_{Q,T} \qquad \text{in } \Omega. \tag{4.2.3c}$$

Here, $L_\Omega(t)$, $t \in [0,T]$, stands for the second order elliptic operator

$$L_\Omega(t) := \frac{1}{2}\sum_{k,\ell=1}^{2} \xi_{k\ell} S_k S_\ell \frac{\partial^2}{\partial S_k \partial S_\ell} + r \sum_{k=1}^{2} S_k \frac{\partial}{\partial S_k} - r. \tag{4.2.4}$$

The boundary functions y_{Σ_3} and y_{Σ_4} represent cash settlements at the lower and at the upper barrier. We have set $y_{\Sigma_3} = 0$. The final time function $y_{Q,T}$ and the boundary function y_{Σ_4} are chosen below. The other two boundary functions y_{Σ_ν}, $1 \le \nu \le 2$, have to be computed as the solutions of the one-dimensional Black-Scholes equations

$$\frac{\partial y_{\Sigma_\nu}}{\partial t} + L_{\Gamma_\nu}(t) y_{\Sigma_\nu} = 0 \qquad \text{in } \Sigma_\nu := \Gamma_\nu \times (0,T) \tag{4.2.5a}$$

$$y_{\Sigma_\nu}(S_\nu, t) = \begin{cases} 0 & \text{if } S_\nu = K_{min} \\ y_{\Sigma_4} & \text{if } S_\nu = K_{max} \end{cases} \quad , \quad t \in (0,T), \tag{4.2.5b}$$

$$y_{\Sigma_\nu}(\cdot, T) = y_{Q,T}|_{\Gamma_\nu} \quad \text{in } \Gamma_\nu \tag{4.2.5c}$$

where $L_{\Gamma_\nu}(t)$, $1 \le \nu \le 2$, $t \in [0,T]$, are the second order elliptic operators

$$L_{\Gamma_\nu}(t) := \frac{1}{2}\sigma_\nu^2 S_\nu^2 \frac{\partial^2}{\partial S_\nu^2} + r S_\nu \frac{\partial}{\partial S_\nu} - r. \tag{4.2.6}$$

Next, we come to the definition of $y_{Q,T}$ and y_{Σ_4}. For that purpose let $M \in \mathbb{N}$ and

$$K =: K_0 < K_1 < \cdots < K_M := K_{max}$$

be a partition of $[K, K_{max}]$ with $K_i := K + i\delta_{|S|}$, $0 \le i \le M$, $\delta_{|S|} := (K_{max} - K)/M$ and setting formally $u := (u_1, \cdots, u_M)^T \in \mathbb{R}_+^M$, as well as $K_{-1} = K_{min}$, $u_{-1} = u_0 = 0$, we may choose $y_4 = u_M$ and $y_T = g(u)$ where

$$(g(u))(S) = u_{i-1} g_1^{(i)}(S) + u_i g_2^{(i)}(S) \qquad \text{for } |S| \in [K_{i-1}, K_i] \text{ and } i = 0, \cdots, M \tag{4.2.7}$$

with

$$g_1^{(i)}(S) := (K_i - S)/\delta_{|S|} \quad \text{and} \quad g_2^{(i)}(S) := (S - K_{i-1})/\delta_{|S|}.$$

We may consider u as a control vector that has to be chosen such that the Greek $\Delta := \nabla y$ per asset point is as close to a prespecified profit $d = (d_1, d_2)^T$ as possible. The controls are subject to the constraints

$$\begin{aligned} u \in & U_{ad} := \{v = (v_1, \cdots, v_M)^T \in \mathbb{R}^M \mid v_i \in U_{ad}^{(i)},\ 1 \le i \le M\}, \\ & U_{ad}^{(i)} := \{v_i \in \mathbb{R} \mid u_{i,min} \le v_i \le u_{i,max}\}. \end{aligned} \tag{4.2.8}$$

We consider the following optimal control problem for the two-dimensional Black-Scholes equation: Find (y_Q, u) such that

$$\inf_{y_Q, u} J(y_Q, u) := \frac{1}{2} \int_0^T \int_\Omega |\nabla y_Q - d|^2 dS dt, \tag{4.2.9}$$

subject to (4.2.3a)-(4.2.3c), (4.2.5a)-(4.2.5c), and (4.2.8).

4.3. Variational formulation of the optimal control problem

We use standard notation from Lebesgue and Sobolev space theory. In particular, given a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$, $d \in \mathbb{N}$, with boundary $\Gamma := \partial\Omega$, for $D \subseteq \Omega$ we refer to $L^p(D)$, $1 \le p \le \infty$ as the Banach spaces of p-th power integrable functions ($p < \infty$) and essentially bounded functions ($p = \infty$) on D with norm $\|\cdot\|_{L^p(D)}$. We denote by $L^p(D)_+$ the positive cone in $L^p(D)$, i.e., $L^p(D)_+ := \{v \in L^p(D) \mid v \ge 0 \text{ a.e. in } D\}$. In case $p = 2$, the space $L^2(D)$ is a Hilbert space whose inner product and norm will be referred to as $(\cdot,\cdot)_{L^2(D)}$.
For $m \in \mathbb{N}_0$ and weight functions $\omega = (\omega_\alpha)_{|\alpha| \le m}$ with $\omega_\alpha \in L^\infty(D)_+$, $\alpha = (\alpha_1, \cdots, \alpha_d) \in \mathbb{N}_0^d$, $|\alpha| := \sum_{i=1}^d \alpha_i$, we denote by $W_\omega^{m,p}(D)$ the weighted Sobolev spaces with norms

$$\|v\|_{W_\omega^{m,p}(D)} := \begin{cases} \Big(\sum\limits_{|\alpha| \le m} \|\omega_\alpha D^\alpha v\|_{L^p(D)}^p \Big)^{1/p} & \text{if } p < \infty \\ \max\limits_{|\alpha| \le m} \|\omega_\alpha D^\alpha v\|_{L^\infty(D)} & \text{if } p = \infty \end{cases}$$

and refer to $|\cdot|_{W_\omega^{m,p}(D)}$ as the associated seminorms. In particular, for $|\alpha| = 1$ we use the notation $\nabla_\omega v := (S_1 \partial v/\partial S_1, \cdots, S_d \partial v/\partial S_d)^T$. For $p < \infty$ and $s \in \mathbb{R}_+$, $s = m + \sigma$, $m \in \mathbb{N}_0$, $0 < \sigma < 1$, we define the weighted Sobolev space $W_\omega^{s,p}(D)$ with norm $\|\cdot\|_{W_\omega^{s,p}(D)}$ in analogy to the standard, non-weighted case and refer to $W_{\omega,0}^{s,p}(D)$ as the closure of $C_0^\infty(D)$ in $W_\omega^{s,p}(D)$. For $s < 0$, we denote by $W_\omega^{-s,p}(D)$ the dual space of $W_{\omega,0}^{-s,q}(D)$, $p^{-1} + q^{-1} = 1$. In case $p = 2$, the spaces $W_\omega^{s,2}(D)$ are Hilbert spaces. We will write $H_\omega^s(D)$ instead of $W_\omega^{s,2}(D)$ and refer to $(\cdot,\cdot)_{H_\omega^s(D)}$ and $\|\cdot\|_{H_\omega^s(D)}$ as the inner products and associated norms. In the standard case $\omega_\alpha \equiv 1$, $|\alpha| \le m$, we will drop the subindex ω.

For a Banach space X and its dual X^*, we refer to $\langle \cdot, \cdot \rangle_{X^*,X}$ as the dual pairing between X^* and X. For Banach spaces X_i, $1 \le i \le n$, $n \in \mathbb{N}$, and a function $v \in \bigcap_{i=1}^n X_i$, we refer to $\|v\|_{\bigcap_{i=1}^n X_i}$ as the norm

$$\|v\|_{\bigcap_{i=1}^n X_i} := \max_{1 \le i \le n} \|v\|_{X_i}. \tag{4.3.1}$$

Moreover, for $T > 0$ and a Banach space X, we denote by $L^p((0,T),X)$, $1 \le p \le \infty$, and $C([0,T],X)$ the Banach spaces of functions $v : [0,T] \to X$ with norms

$$\|v\|_{L^p((0,T),X)} := \begin{cases} \Big(\int\limits_0^T \|v(t)\|_X^p dt \Big)^{1/p} & \text{if } 1 \le p < \infty \\ \operatorname*{ess\,sup}_{t \in [0,T]} \|v(t)\|_X & \text{if } p = \infty \end{cases}$$
$$\|v\|_{C([0,T],X)} : = \max_{t \in [0,T]} \|v(t)\|_X.$$

The spaces $W^{s,p}((0,T),X)$ and $H^s((0,T),X)$, $s \in \mathbb{R}_+$, are defined likewise.
In particular, for a subspace $V \subset H^1_\omega(\Omega)$ with dual V^* we will consider the space

$$H^1((0,T),V^*) \cap L^2((0,T),V), \tag{4.3.2}$$

and note that the following continuous embedding holds true

$$H^1((0,T),V^*) \cap L^2((0,T),V) \subset C([0,T],L^2(\Omega)). \tag{4.3.3}$$

For $y \in H^1((0,T),V^*) \cap L^2((0,T),V)$, we further denote by $\gamma_{\Sigma'}(y)$, $\Sigma' \subset \Sigma := \Gamma \times (0,T)$, the trace of y on Σ'.
In the sequel, $\Omega \subset \mathbb{R}^2_+$ will stand for the trapezoidal domain given by (4.2.1) and $\Gamma_i \subset \partial\Omega$, $1 \le i \le 4$, for its boundaries as specified by (4.2.2). The weight functions ω_α will be chosen according to

$$w_\alpha = S_{\alpha_1} S_{\alpha_2} \ , \ \alpha = (\alpha_1, \alpha_2) \ , \ |\alpha| \le 2, \tag{4.3.4}$$

with the convention that $S_{\alpha_i} = 1$ for $\alpha_i = 0$, $1 \le i \le 2$.

We reformulate the boundary and final time problems for the backward parabolic equations as initial-boundary value problems by means of the transformation $t \mapsto T - t$. Keeping for notational simplicity the same notation for y_Q and y_{Σ_ν} and the operators L_Ω, L_{Σ_ν}, $1 \le \nu \le 2$, the initial-boundary value problems read as follows:

$$\frac{\partial y_Q}{\partial t} - L_\Omega(t) y_Q = 0 \quad \text{in } Q, \tag{4.3.5a}$$

$$y_Q = \begin{cases} y_{\Sigma_\nu} & \text{on } \Sigma_\nu, \ 1 \le \nu \le 2, \\ 0 & \text{on } \Sigma_3, \\ u_M & \text{on } \Sigma_4, \end{cases} \tag{4.3.5b}$$

$$y_Q(\cdot, 0) = g(u) \quad \text{in } \Omega, \tag{4.3.5c}$$

$$\frac{\partial y_{\Sigma_\nu}}{\partial t} - L_{\Gamma_\nu}(t) y_{\Sigma_\nu} = 0 \quad \text{in } \Sigma_\nu, \tag{4.3.6a}$$

$$y_{\Sigma_\nu}(S_\nu, t) = \begin{cases} 0 & \text{if } S_\nu = K_{min} \\ u_M & \text{if } S_\nu = K_{max} \end{cases}, \; t \in (0,T), \tag{4.3.6b}$$

$$y_{\Sigma_\nu}(\cdot, 0) = g(u)_{|\Gamma_\nu}. \tag{4.3.6c}$$

As far as the volatilities σ_k, $1 \leq k \leq 2$, and the interest r are concerned, we will impose the following assumptions:

(A1) The volatilities satisfy $\sigma_k \in C([0,T], C^2(\overline{\Omega}))$, $1 \leq k \leq 2$, and there exist constants $\sigma_k^{(min)} > 0$, $C_{\sigma_k} > 0$, such that

$$\sigma_k(S,t) \geq \sigma_k^{(min)}, \qquad (S,t) \in \bar{Q}, \; 1 \leq k \leq 2, \tag{4.3.7a}$$

$$|S \cdot \nabla \sigma_k(S,t)| \leq C_{\sigma_k} \qquad (S,t) \in \bar{Q}, \; 1 \leq k \leq 2. \tag{4.3.7b}$$

(A2) The interest rate satisfies $r \in C([0,T])$ such that $r(t) > 0$, $t \in [0,T]$.

For the correlation matrix $\xi = (\xi)_{k,\ell=1}^2$, it is an immediate consequence of assumption **(A1)** that $\xi_{k,\ell} \in C([0,T], C^2(\overline{\Omega}))$, $1 \leq k, \ell \leq 2$, and that there exists a constant $\xi_{min} > 0$ such that for all $\eta \in \mathbb{R}^2$ there holds

$$\sum_{k,\ell=1}^{2} \xi_{k,\ell}(S,t) \eta_k \eta_\ell \geq \xi_{min} |\eta|^2, \quad (S,t) \in \bar{Q}. \tag{4.3.8}$$

We now study the weak formulations of the initial-boundary value problems (4.3.5a)-(4.3.5c) and (4.3.6a)-(4.3.6c):

A function $y_Q \in H^1((0,T), V^*) \cap L^2((0,T), V)$, where $V := \{v \in H^1_\omega(\Omega) \mid v|_{\Sigma_\nu} = y_{\Sigma_\nu}, 1 \leq \nu \leq 2, v|_{\Sigma_3} = 0, v|_{\Sigma_4} = u_M\}$, is called a weak solution of (4.3.5a)-(4.3.5c), if for all $v \in L^2((0,T), H^1_{\omega,0}(\Omega))$ there holds

$$\int_0^T \langle \frac{\partial y_Q}{\partial t}, v \rangle_{H^{-1}_\omega(\Omega), H^1_{\omega,0}(\Omega)} dt + \int_0^T a(t; y_Q, v) dt = 0, \tag{4.3.9a}$$

$$y_Q(\cdot, 0) = g(u) \tag{4.3.9b}$$

Here, the bilinear form $a(t; \cdot, \cdot)$, $t \in (0,T)$, is given by

$$a(t; y, v) := \int_\Omega \Bigg(\frac{1}{2} \sum_{k,\ell=1}^{2} \xi_{k\ell} S_k S_\ell \frac{\partial y}{\partial S_k} \frac{\partial v}{\partial S_\ell} + \sum_{k=1}^{2} \left(\frac{1}{2} \sum_{\ell=1}^{2} \frac{\partial \xi_{k\ell}}{\partial S_\ell} S_\ell + \xi_{kk} + \frac{1}{2}\xi_{12} - r \right) S_k \frac{\partial y}{\partial S_k} v + ryv \Bigg) dS.$$

Likewise, a function $y_{\Sigma_\nu} \in H^1((0,T),V_\nu^*) \cap L^2((0,T),V_\nu)$, where $V_\nu := \{v \in H^1_\omega(\Gamma_\nu) \mid v(K_{min}) = 0,\ v(K_{max}) = u_M\}$, is said to be a weak solution of (4.3.6a)-(4.3.6c), if for all $v_\nu \in L^2((0,T),H^1_{\omega,0}(\Gamma_\nu))$ there holds

$$\int_0^T \langle \frac{\partial y_{\Sigma_\nu}}{\partial t}, v_\nu\rangle_{H_\omega^{-1}(\Gamma_\nu),H^1_{\omega,0}(\Gamma_\nu)} dt + \int_0^T a_\nu(t; y_{\Sigma_\nu}, v_\nu) dt = 0, \tag{4.3.10a}$$

$$y_{\Sigma_\nu}(\cdot,0) = g(u)|_{\Gamma_\nu}. \tag{4.3.10b}$$

Here, the bilinear form $a_\nu(t;\cdot,\cdot)$, $t \in (0,T)$, is given by

$$a_\nu(t;y,v) := \int_{\Gamma_\nu} \left(\frac{1}{2}\sigma_\nu^2 S_\nu^2 \frac{\partial y}{\partial S_\nu}\frac{\partial v}{\partial S_\nu} + \left(\sigma_\nu \frac{\partial \sigma_\nu}{\partial S_\nu} S_\nu + \sigma_\nu^2 - r\right) S_\nu \frac{\partial y}{\partial S_\nu} v + ryv\right) dS_\nu.$$

Theorem 4.3.1. *For arbitrarily chosen but fixed $u \in \mathbb{R}^M$, the state equations (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b) admit unique solutions $y_Q \in C([0,T],V) \cap L^2((0,T),V \cap H^2_\omega(\Omega))$ and $y_{\Sigma_\nu} \in C([0,T],V_\nu) \cap L^2((0,T),V_\nu \cap H^2_\omega(\Gamma_\nu))$, $1 \le \nu \le 2$. Moreover, for all $t \in (0,T)$ there holds*

$$exp(-2\lambda t)\|y_Q(t)\|^2_{L^2(\Omega)} + 2\xi^2_{min} \int_0^t exp(-2\lambda\tau)|y_Q(\tau)|^2_V \ d\tau \le \|g(u)\|^2_{L^2(\Omega)}, \tag{4.3.11}$$

$$exp(-2\lambda_\nu t)\|y_{\Sigma_\nu}(t)\|^2_{L^2(\Gamma_\nu)} + \frac{1}{2}\ (\sigma_\nu^{(min)})^2 \int_0^t exp(-2\lambda_\nu\tau)|y_{\Sigma_\nu}(\tau)|^2_{V_\nu}\ d\tau \le \|g(u)\|^2_{L^2(\Gamma_\nu)},$$
$$1 \le \nu \le 2.$$

Proof. It follows from assumptions (A_1), (A_2) and the Poincaré inequalities (cf., e.g., [AP05])

$$\|v\|^2_{L^2(\Omega)} \le 2 \int_\Omega |S \cdot \nabla_\omega v(S)|^2 \ dS_1 dS_2,\ v \in V,$$
$$\|v_\nu\|^2_{L^2(\Gamma_\nu)} \le 2 \int_{\Gamma_\nu} |S_\nu \frac{\partial v_\nu}{\partial S_\nu}(S_\nu)|^2 \ dS_\nu,\ v_\nu \in V_\nu,\ 1 \le \nu \le 2,$$

that the bilinear forms $a(t;\cdot,\cdot)$ and $a_\nu(t;\cdot,\cdot)$ satisfy Gârding-type inequalities uniformly in t, i.e., there exist constants $\lambda \ge 0$ and $\lambda_\nu \ge 0$, $1 \le \nu \le 2$, such that for all $t \in (0,T)$ there holds

$$a(t;v,v) \ge \frac{1}{4}\ \xi^2_{min}\ \|v\|^2_V - \lambda\|v\|^2_{L^2(\Omega)},\ v \in V,$$
$$a_\nu(t;v,v) \ge \frac{1}{4}\ (\sigma_\nu^{(min)})^2\ \|v\|^2_{V_\nu} - \lambda_\nu\|v\|^2_{L^2(\Gamma_\nu)},\ v \in V_\nu,\ 1 \le \nu \le 2.$$

Consequently, the initial-boundary value problems (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b) have unique solutions $y_Q \in H^1((0,T),V^*) \cap L^2((0,T),V)$ and $y_{\Sigma_\nu} \in H^1((0,T),V_\nu^*) \cap L^2((0,T);V_\nu)$, $1 \le \nu \le 2$, satisfying (4.3.11) (cf., e.g., Thm. 2.11 and section 2.6 in [AP05]). Moreover, standard regularity results for parabolic partial differential equations [AP05, RR93] reveal $y_Q \in C([0,T],V) \cap L^2((0,T),V \cap H^2_\omega(\Omega))$ and $y_{\Sigma_\nu} \in C([0,T],V_\nu) \cap L^2((0,T),V_\nu \cap H^2_\omega(\Gamma_\nu))$, $1 \le \nu \le 2$. □

Based on the weak formulation of the state equations, the optimal control problem from section 4.2 reads as follows:
Find (y,u), where $y = (y_Q, y_{\Sigma_1}, y_{\Sigma_2})$, $y_Q \in H^1((0,T),V^*) \cap L^2((0,T),V)$, $y_{\Sigma_\nu} \in H^1((0,T),V_\nu^*) \cap L^2((0,T);V_\nu)$, $1 \le \nu \le 2$, and $u \in U_{ad}$ such that

$$\inf_{y,u} J(y,u) := \frac{1}{2}\int_0^T \int_\Omega |\nabla y_Q - d|^2 dS dt, \tag{4.3.12a}$$

$$\text{subject to (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b).} \tag{4.3.12b}$$

Theorem 4.3.2. *The optimal control problem* (4.3.12a), (4.3.12b) *admits a unique solution* (y,u).

Proof. We denote by $S : U_{ad} \to H^1((0,T),V^*) \cap L^2((0,T),V)$ and $S_\nu : U_{ad} \to H^1((0,T),V_\nu^*) \cap L^2((0,T);V_\nu)$, $1 \le \nu \le 2$, the control-to-state maps which assign to an admissible control $u \in U_{ad}$ the unique solutions y_Q and y_{Σ_ν}, $1 \le \nu \le 2$, of the state equations (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b). Replacing $y = (y_Q, y_{\Sigma_1}, y_{\Sigma_2})$ in (4.3.12a) with $(S(u), S_1(u), S_2(u))$, the reduced formulation of (4.3.12a), (4.3.12b) is given by:

$$\inf_{u \in U_{ad}} J_{red}(u) := \frac{1}{2}\int_0^T \int_\Omega |\nabla S(u) - d|^2 dS dt, \tag{4.3.13a}$$

such that the triple $(S(u), S_1(u), S_2(u))$

$$\text{satisfies (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b).} \tag{4.3.13b}$$

Let $(u_n)_{n\in\mathbb{N}}$, $u_n \in U_{ad}$, $n \in \mathbb{N}$, be a minimizing sequence. Due to the facts that $(u_n)_{n\in\mathbb{N}}$ is bounded and $U_{ad} \subset \mathbb{R}^M$ is closed, there exist a subsequence $\mathbb{N}' \subset \mathbb{N}$ and $u^* \in U_{ad}$ such that $u_n \to u^* (\mathbb{N}' \ni n \to \infty)$. From the continuity of the control-to-state maps we deduce

$$S(u_n) \to S(u^*), \quad S_\nu(u_n) \to S_\nu(u^*) \quad (\mathbb{N}' \ni n \to \infty).$$

Moreover, $(S(u^*), S_1(u^*), S_2(u^*))$ satisfies (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b). Taking additionally the continuity of g into account, we find

$$J_{red}(u_n) \to J_{red}(u^*) \quad (\mathbb{N}' \ni n \to \infty),$$

which allows to conclude. □

4.4. Necessary optimality conditions

The first order necessary optimality conditions can be stated in terms of (y, u), $y = (y_Q, y_{\Sigma_1}, y_{\Sigma_2})$, and an adjoint state $p \in W_0(0,T)$ that is the solution of a final time problem on Q with homogeneous Dirichlet boundary conditions.

Theorem 4.4.1. *If $(y, u) \in W(0,T) \times U_{ad}$ is the optimal solution of (4.3.12a), (4.3.12b), there exists $p_Q \in W_0(0,T)$ such that there holds:*

(i) p_Q is the weak solution of the parabolic final time problem

$$-\frac{\partial p_Q}{\partial t} - A^* p_Q = -\nabla \cdot (\nabla y_Q - d) \quad \text{in } Q, \tag{4.4.1a}$$

$$p_Q = 0 \quad \text{on } \Sigma, \tag{4.4.1b}$$

$$p_Q(\cdot, T) = 0 \quad \text{in } \Omega. \tag{4.4.1c}$$

(ii) The variational inequality

$$\left(\int_0^T \left(\gamma_{\Gamma_4} \left(\eta_{\Gamma_4} \cdot R_{\Gamma_4}(p_Q) \right) \right) dt - g_u^*(u) p_Q(0) \right) \cdot (v - u) \geq 0, \;\; v \in U_{ad} \tag{4.4.2}$$

is satisfied, where $R_{\Sigma_4}(p)$ is given by

$$R_{\Gamma_4}(p_Q) = \left(R_{\Gamma_4}^{(1)}(p_Q), R_{\Gamma_4}^{(2)}(p_Q) \right)^T, \tag{4.4.3}$$

$$R_{\Gamma_4}^{(k)}(p_Q) := S_k \left(\frac{1}{2} \sum_{\ell=1}^{2} \xi_{k\ell} S_\ell \frac{\partial p_Q}{\partial S_\ell} - r p_Q \right), \;\; 1 \leq k \leq 2,$$

and $g_u^*(u) \in \mathcal{L}(L^2(\Omega), \mathbb{R}^M)$ stands for the adjoint of the Fréchet derivative of g at $u \in U_{ad}$.

Proof. We introduce multipliers

$$p = \left(p_Q, (p_{\Sigma_\nu})_{\nu=1}^2 \right),$$

$$q = \left((q_{\Sigma_\nu})_{\nu=1}^4, (q_{K_{max},\nu})_{\nu=1}^2, \left(q_{K_{min},\nu} \right)_{\nu=1}^2, q_{0,\Omega}, (q_{0,0,\Gamma_\nu})_{\nu=1}^2 \right),$$

such that

$$p_Q \in W_0(0,T), \;\; p_{\Sigma_\nu} \in W_{\nu,0}(0,T), \;\; 1 \leq \nu \leq 2$$

$$q_{\Sigma_\nu} \in L^2\left((0,T), H_\omega^{-1/2}(\Gamma_\nu) \right), \;\; 1 \leq \nu \leq 4, \;\; q_{K_{max},\nu}, q_{K_{min},\nu} \in \mathbb{R}, \;\; 1 \leq \nu \leq 2,$$

$$q_{0,\Omega} \in L^2(\Omega), \;\; q_{0,0,\Gamma_\nu} \in L^2(\Gamma_\nu), \;\; 1 \leq \nu \leq 2.$$

We consider the Lagrangian

$$
\begin{aligned}
&L(y,u,p,q) := \\
&\quad J(y_Q,u) + \int_0^T \langle \frac{\partial y_Q}{\partial t} - A(t)y_Q, p_Q\rangle dt + \sum_{\nu=1}^{2} \int_0^T \langle \frac{\partial y_{\Sigma_\nu}}{\partial t} - A_\nu(t)y_{\Sigma_\nu}, p_{\Sigma_\nu}\rangle dt \\
&\quad + \sum_{\nu=1}^{4} \int_0^T \langle q_{\Sigma_\nu}, y_{\Sigma_\nu} - \gamma_{\Sigma_\nu}(y_Q)\rangle dt \\
&\quad + \sum_{\nu=1}^{2} \int_0^T \Big(q_{K_{max},\nu}\left(u_M - \gamma_{K_{max},\nu}(y_{\Sigma_\nu})\right) - q_{K_{min},\nu}(y_{\Sigma_\nu})\Big) \\
&\quad + \left(y_Q(0) - g(u), q_{0,\Omega}\right)_{L^2(\Omega)} + \sum_{\nu=1}^{2} \left(y_{\Sigma_\nu}(0) - g(u), q_{0,0,\Gamma_\nu}\right)_{L^2(\Gamma_\nu)},
\end{aligned}
$$

where $\gamma_{\Sigma_\nu}(y_Q)$, $1 \le \nu \le 2$, is the trace of y_Q on Σ_ν and

$$\gamma_{K,\nu}(y_{\Sigma_\nu}) = y_{\Sigma_\nu}(K,\cdot),\ K \in \{K_{min}, K_{max}\},\ 1 \le \nu \le 2.$$

Denoting by $A_\nu^*(t)$ the adjoint of $A_\nu(t)$, $1 \le \nu \le 2$, and introducing

$$R_K(p_{\Sigma_\nu}) := \frac{1}{2}\sigma_\nu^2 S_K^2 \frac{\partial p_{\Sigma_\nu}}{\partial S_\nu}(K) - rS_K p_{\Sigma_\nu}(K),\ 1 \le \nu \le 2,\ K \in \{K_{min}, K_{max}\},$$

integration by parts yields in time and an application of Green's formula results in

$$
\begin{aligned}
&L(y,u,p,q) := \qquad\qquad (4.4.4) \\
&\quad J(y_Q,u) + \int_0^T \langle -\frac{\partial p_Q}{\partial t} - A^*(t)p_Q, y_Q\rangle dt + \sum_{\nu=1}^{2} \int_0^T \langle \frac{\partial y_{\Sigma_\nu}}{\partial t} - A_\nu(t)y_{\Sigma_\nu}, p_{\Sigma_\nu}\rangle dt \\
&\quad + \sum_{\nu=1}^{4} \int_0^T \left(\langle \gamma_{\Sigma_\nu}\left(\eta_{\Sigma_\nu} \cdot R_{\Sigma_\nu}(p_Q)\right) - q_{\Sigma_\nu}, \gamma_{\Sigma_\nu}(y_Q)\rangle + \langle q_{\Sigma_\nu}, y_{\Sigma_\nu}\rangle\right) dt \\
&\quad + \sum_{\nu=1}^{2} \int_0^T \left(\left(R_{K_{max}}(p_{\Sigma_\nu}) - q_{K_{max},\nu}\right)\gamma_{K_{max},\nu}(y_{\Sigma_\nu}) + q_{K_{max},\nu} u_M\right) dt \\
&\quad - \sum_{\nu=1}^{2} \int_0^T \Big(\Big(R_{K_{min}}(p_{\Sigma_\nu}) - q_{K_{min},\nu}\Big)\gamma_{K_{min},\nu}(y_{\Sigma_\nu})\Big) dt \\
&\quad + \left(y_Q(0), q_{0,\Omega} - p_Q(0)\right)_{L^2(\Omega)} - \left(g(u), q_{0,\Omega}\right)_{L^2(\Omega)} + \left(y_Q(T), p_Q(T)\right)_{L^2(\Omega)} \\
&\quad + \sum_{\nu=1}^{2} \Big(\left(y_{\Sigma_\nu}, q_{0,\Gamma_\nu} - p_{\Sigma_\nu}(0)\right)_{L^2(\Gamma_\nu} - \left(g(u), q_{0,\Gamma_\nu}\right)_{L^2(\Gamma_\nu} + \left(y_{\Sigma_\nu}, p_{\Sigma_\nu}(T)\right)_{L^2(\Gamma_\nu}\Big).
\end{aligned}
$$

Here, $R_{\Sigma_\nu}(p_Q)$, $1 \leq \nu \leq 3$, is defined as in (4.4.3) with Σ_4 replaced by Σ_ν, $1 \leq \nu \leq 3$. In view of $J_y(y_Q, u) = -\nabla \cdot (\nabla y_Q - d)$, the optimality conditions

$$L_y(y,u,p,q) = 0 \quad \text{and} \quad L_q(y,u,p,q) = 0$$

reveals that

$$\begin{aligned}
q_{\Sigma_\nu} &= \gamma_{\Sigma_\nu}(n_{\Sigma_\nu} \cdot R_{\Sigma_\nu}(p_Q)), & & & 1 \leq \nu \leq 4, \quad &(4.4.5a)\\
q_{K_{min},\nu} &= R_{K_{min}}(p_{\Sigma_\nu}), & q_{K_{max},\nu} &= R_{K_{max}}(p_{\Sigma_\nu}), & 1 \leq \nu \leq 2, \quad &(4.4.5b)\\
q_{0,\Omega} &= \gamma_{0,\Omega}(p_Q), & q_{0,\Gamma_\nu} &= p_{\Sigma_\nu}(0), & 1 \leq \nu \leq 2, \quad &(4.4.5c)
\end{aligned}$$

and

$$\begin{aligned}
\gamma_{\Sigma_\nu}(y_Q) &= y_{\Sigma_\nu}, & & & 1 \leq \nu \leq 4, \quad &(4.4.6a)\\
\gamma_{K_{min},\nu}(y_{\Sigma_\nu}) &= 0, & \gamma_{K_{max},\nu}(y_{\Sigma_\nu}) &= u_M, & 1 \leq \nu \leq 2, \quad &(4.4.6b)\\
y_Q(\cdot,0) &= g(u), & y_{\Sigma_\nu}(\cdot,0) &= g(u)|_{\Sigma_\nu}, & 1 \leq \nu \leq 2. \quad &(4.4.6c)
\end{aligned}$$

Further, p_{Σ_ν}, $1 \leq \nu \leq 2$, is the weak solution of

$$\begin{aligned}
-\frac{\partial p_{\Sigma_\nu}}{\partial t} - A_\nu^*(t) p_{\Sigma_\nu} &= n_{\Sigma_\nu} \cdot R_{\Sigma_\nu}(p_Q) && \text{in } \Sigma_\nu,\\
R_K(p_{\Sigma_\nu}) &= 0, && K \in \{K_{min}, K_{max}\},\\
p_{\Sigma_\nu}(\cdot, T) &= 0.
\end{aligned}$$

Since $n_{\Sigma_\nu} \cdot R_{\Sigma_\nu}(p_Q) = 0$, $1 \leq \nu \leq 2$, it follows that $p_{\Sigma_\nu} = 0$ and hence $p_{\Sigma_\nu} = \gamma_{\Sigma_\nu}(p_Q)$, $1 \leq \nu \leq 2$.
Taking (4.4.5a)-(4.4.5c) and (4.4.6a)-(4.4.6c) into account, the optimality condition

$$L_p(y,u,p,q) = 0$$

shows that y_Q and y_{Σ_ν}, $1 \leq \nu \leq 2$, are the weak solutions of (4.3.5a)-(4.3.5c) and (4.3.6a)-(4.3.6c). Finally, observing (4.2.7) and $y_{\Sigma_4} = u_M$ as well as the regularity results of Theorem 4.3.1, the optimality condition

$$(L_u(y,u,p,q)) \cdot (v - u) \geq 0 \quad , \quad v \in U_{ad},$$

gives rise to (4.4.2). □

4.5. Semi-discretization of the optimal control problem

The parabolic problems (4.3.9a), (4.3.9b) and (4.3.10a), (4.3.10b) will be discretized in space by conforming P1 finite elements. To this end, we consider a shape-regular simplicial triangulation $\mathcal{T}_h(\Omega)$ of Ω which aligns with Γ_j, $1 \leq j \leq 4$, so that this triangulation generates triangulations $\mathcal{T}_h(\Gamma_j)$ of Γ_j, $1 \leq j \leq 4$, as well. We refer to $\mathcal{N}_h(D)$ and $\mathcal{E}_h(D)$, $D \subseteq \overline{\Omega}$, as the sets of vertices in $D \subseteq \overline{\Omega}$. We denote by h_T

and $|T|$ the diameter and area of an element $T \in \mathcal{T}_h^{(m)}(\Omega)$. For $D \subset \overline{\Omega}$, we refer to $P_k(D)$, $k \in \mathbb{N}_0$, as the linear spaces of polynomials of degree $\leq k$ on D.
We define V_h as the finite element space of continuous P1 finite elements associated with the triangulation $\mathcal{T}_h(\Omega)$, i.e.,

$$V_h := \{v_h \in C(\bar{\Omega}) \mid v_h|_K \in P_1(K),\ K \in \mathcal{T}_h(\Omega)\}, \tag{4.5.1}$$

and we set $V_{h,0} := V_h \cap C_0(\bar{\Omega})$. Likewise, we define $V_{h,\nu}$, $1 \leq \nu \leq 2$, as the finite element spaces of continuous P1 finite elements associated with the triangulations $\mathcal{T}_h(\Gamma_\nu)$ attaining the values 0 at $S_\nu = K_{min}$ and u_M at $S_\nu = K_{max}$, i.e.,

$$\begin{aligned} V_{h,\nu} := \{v_h \in C(\bar{\Gamma}_\nu) \mid v_h|_K \in P_1(K),\ K \in \mathcal{T}_h(\Gamma_\nu)\ , \\ v_h(K_{min}) = 0\ ,\ v_h(K_{max}) = u_M\}, \end{aligned} \tag{4.5.2}$$

and we define $V_{h,\nu,0}$ in the same way, but replacing u_M with 0.
The semi-discrete approximation of (4.3.9a), (4.3.9b) amounts to the computation of $y_{h,Q} \in C^1([0,T], V_h)$ with $y_{h,Q}(\cdot,t)|_{\Gamma_\nu} = y_{h,\Gamma_\nu}(\cdot,t)$, $1 \leq \nu \leq 2$, and $y_{h,Q}(\cdot,t)|_{\Gamma_3} = 0$, $y_{h,Q}(\cdot,t)|_{\Gamma_4} = u_M$, such that

$$(\frac{dy_{h,Q}}{dt}, v_h)_{L^2(\Omega)} + a(t; y_{h,Q}, v_h) = 0, \qquad v_h \in V_{h,0}, \tag{4.5.3a}$$

$$(y_{h,Q}(\cdot,0), v_h)_{L^2(\Omega)} = (g(u), v_h)_{L^2(\Omega)}, \quad v_h \in V_h. \tag{4.5.3b}$$

On the other hand, for the semi-discrete approximation of (4.3.10a), (4.3.10b) we have to compute $y_{h,\Gamma_\nu} \in C^1([0,T], V_{h,\nu})$, $1 \leq \nu \leq 2$, such that

$$(\frac{dy_{h,\Gamma_\nu}}{dt}, v_h)_{L^2(\Gamma_\nu)} + a_\nu(t; y_{h,\Gamma_\nu}, v_h) = 0, \qquad v_h \in V_{h,\nu,0}, \tag{4.5.4a}$$

$$(y_{h,\Gamma_\nu}(\cdot,0), v_h)_{L^2(\Gamma_\nu)} = (g(u), v_h)_{L^2(\Gamma_\nu)} \quad v_h \in V_{h,\nu}. \tag{4.5.4b}$$

The semi-discrete optimal control problems reads: Find (y_h, u), where $y_h = (y_{h,Q}, y_{h,\Gamma_1}, y_{h,\Gamma_2})$ such that

$$\inf_{y_h,u} J_h(y_h,u) := \frac{1}{2} \int_0^T \sum_{K \in \mathcal{T}_h(\Omega)} \|\nabla y_{h,Q}(\cdot,t) - d\|^2_{L^2(K)}\, dt, \tag{4.5.5a}$$

$$\text{subject to (4.5.4a), (4.5.4b), (4.5.3a), (4.5.3b) and (4.2.8).} \tag{4.5.5b}$$

The semi-discrete problem (4.5.5a), (4.5.5b) will be written in a more compact form as a control constrained optimal control problem for an initial-value problem associated with a system of first order linear ordinary differential equations. To this end, we set

$$N_\Omega := \text{card}(\mathcal{N}_h(\Omega)),\ N_{\Gamma_\nu} := \text{card}(\mathcal{N}_h(\Gamma_\nu)),\ 1 \leq \nu \leq 4,\ N := N_\Omega + N_{\Gamma_1} + N_{\Gamma_2},$$

and denote by $\varphi_{h,\Omega}^{(i)}$, $1 \le i \le N_\Omega$, $\varphi_{h,\Gamma_\nu}^{(i)}$, $1 \le i \le N_{\Gamma_\nu}$, the nodal basis functions associated with the nodal points in $\mathcal{N}_h(\Omega)$ and $\mathcal{N}_h(\Gamma_\nu)$, $\nu \in \{1,2,4\}$. We introduce

$$M = \begin{pmatrix} M_\Omega & M_{\Omega\Gamma_1} & M_{\Omega\Gamma_2} \\ 0 & M_{\Gamma_1} & 0 \\ 0 & 0 & M_{\Gamma_2} \end{pmatrix} \quad , \quad A(t) = \begin{pmatrix} A_\Omega(t) & A_{\Omega\Gamma_1}(t) & A_{\Omega\Gamma_2}(t) \\ 0 & A_{\Gamma_1}(t) & 0 \\ 0 & 0 & A_{\Gamma_2}(t) \end{pmatrix} ,$$
$$t \in (0,T],$$

as the block mass matrix $M \in \mathbb{R}^{N\times N}$ and the block stiffness matrix $A \in \mathbb{R}^{N\times N}$. Here, $M_\Omega \in \mathbb{R}^{N_\Omega \times N_\Omega}$, $M_{\Omega\Gamma_\nu} \in \mathbb{R}^{N_\Omega \times N_{\Gamma_\nu}}$ and $M_{\Gamma_\nu} \in \mathbb{R}^{N_{\Gamma_\nu}\times N_{\Gamma_\nu}}$ stand for the matrices

$$\begin{aligned}
(M_\Omega)_{ij} :=& \ (\varphi_{h,\Omega}^{(j)}, \varphi_{h,\Omega}^{(i)})_{L^2(\Omega)} \ , \quad 1 \le i,j \le N_\Omega, \\
(M_{\Omega\Gamma_\nu})_{ij} :=& \ (\varphi_{h,\Gamma_\nu}^{(j)}, \varphi_{h,\Omega}^{(i)})_{L^2(\Omega)} \ , \quad 1 \le i \le N_\Omega \ , \quad 1 \le j \le N_{\Gamma_\nu} \ , \ 1 \le \nu \le 2, \\
(M_{\Gamma_\nu})_{ij} :=& \ (\varphi_{h,\Gamma_\nu}^{(j)}, \varphi_{h,\Gamma_\nu}^{(i)})_{L^2(\Gamma_\nu)} \ , \quad 1 \le i,j \le N_{\Gamma_\nu} \ , \ 1 \le \nu \le 2,
\end{aligned}$$

whereas the matrices $A_\Omega(t) \in \mathbb{R}^{N_\Omega\times N_\Omega}$, $A_{\Omega\Gamma_\nu}(t) \in \mathbb{R}^{N_\Omega \times N_{\Gamma_\nu}}$, $A_{\Gamma_\nu}(t) \in \mathbb{R}^{N_{\Gamma_\nu}\times N_{\Gamma_\nu}}$ are given by

$$\begin{aligned}
(A_\Omega(t))_{ij} :=& \ a(t; \varphi_{h,\Omega}^{(j)}, \varphi_{h,\Omega}^{(i)}) \ , \quad 1 \le i,j \le N_\Omega, \\
(A_{\Omega\Gamma_\nu}(t))_{ij} :=& \ a(t; \varphi_{h,\Gamma_\nu}^{(j)}, \varphi_{h,\Omega}^{(i)}) \ , \quad 1 \le i \le N_\Omega \ , \quad 1 \le j \le N_{\Gamma_\nu} \ , \ 1 \le \nu \le 2, \\
(A_{\Gamma_\nu}(t))_{ij} :=& \ a(t; \varphi_{h,\Gamma_\nu}^{(j)}, \varphi_{h,\Gamma_\nu}^{(i)}) \ , \quad 1 \le i,j \le N_{\Gamma_\nu} \ , \ 1 \le \nu \le 2.
\end{aligned}$$

We further introduce the input matrices

$$B(t) = (0 \ B_M(t)) \quad , \quad B_M(t) = (B_{M,\Omega}(t), B_{M,\Gamma_1}(t), B_{M,\Gamma_2}(t))^T, \ t \in (0,T].$$

Here, $0 \in \mathbb{R}^{N\times(M-1)}$ and $B_{M,\Omega}(t) \in \mathbb{R}^{N_Q\times 1}$ and $B_{M,\Gamma_\nu}(t) \in \mathbb{R}^{N_{\Gamma_\nu}\times 1}$ are defined by

$$\begin{aligned}
(B_{M,\Omega}(t))_i :=& -\sum_{j=1}^{N_{\Gamma_4}} a(t; \varphi_{h,\Gamma_4}^{(j)}, \varphi_{h,\Omega}^{(i)}) \ , \quad 1 \le i \le N_\Omega \\
(B_{M,\Gamma_\nu}(t))_i :=& -a(t; \varphi_{h,\Gamma_4}^{N_{\Gamma_4}^{(\nu)}}, \varphi_{h,\Gamma_\nu}^{(i)}) \ , \qquad 1 \le i \le N_{\Gamma_\nu} \ , \ 1 \le \nu \le 2
\end{aligned}$$

where $N_{\Gamma_4}^{(\nu)} := (2-\nu)+(\nu-1)N_{\Gamma_4}$, $1 \le \nu \le 2$. Let $G \in \mathbb{R}^{N\times M}$, $G = (G_\Omega, G_{\Gamma_1}, G_{\Gamma_2})^T$, $G_\Omega \in \mathbb{R}^{N_\Omega \times M}$ and $G_{\Gamma_\nu} \in \mathbb{R}^{N_{\Gamma_\nu}\times M}$ be defined by

$$(G_\Omega)_{ij} := \int_{\Omega_j} g_2^{(j)}(S)\varphi_{h,\Omega}^{(i)}(S)dS + \int_{\Omega_{j+1}} g_1^{(j+1)}(S)\varphi_{h,\Omega}^{(i)}(S)dS,$$

$$(G_{\Gamma_\nu})_{ij} := \int_{K_{j-1}}^{K_j} g_2^{(j)}(S_\nu)\varphi_{h,\Gamma_\nu}^{(i)}(S_\nu)dS_\nu + \int_{K_j}^{K_{j+1}} g_1^{(j+1)}(S_\nu)\varphi_{h,\Gamma_\nu}^{(i)}(S_\nu)dS_\nu \ , \ 1 \le \nu \le 2$$

and let $C_\Omega \in \mathbb{R}^{N_\Omega \times N_\Omega}$ and $D_\Omega^{(k)} \in \mathbb{R}^{N_\Omega \times N_\Omega}$ be defined by

$$(C_\Omega)_{ij} := \sum_{T \in \mathcal{T}_h(\Omega)} \int_T \nabla \varphi_{h,\Omega}^{(j)} \cdot \nabla \varphi_{h,\Omega}^{(i)} dS \ , \quad 1 \le i,j \le N_\Omega,$$

$$(D_\Omega^{(k)})_{ij} := \sum_{T \in \mathcal{T}_h(\Omega)} \int_T \frac{\partial \varphi_{h,\Omega}^{(j)}}{\partial S_k} \varphi_{h,\Omega}^{(i)} dS \ , \qquad 1 \le i,j \le N_\Omega \ , \ 1 \le k \le 2.$$

The semi-discrete optimal control problem reads as follows: Find $y \in C^1([0,T], \mathbb{R}^N)$, $y = (y_Q, y_{\Sigma_1}, y_{\Sigma_2})$, $u \in U_{ad}$, such that

$$\inf_{y,u} J(y,u) := \frac{1}{2} \int_0^T \Big(y_Q^T C_\Omega y_Q - 2 \sum_{k=1}^2 d_k^T D_\Omega^{(k)} y_Q + \sum_{k=1}^2 d_k^T M_\Omega d_k \Big) \ dt, \tag{4.5.6a}$$

subject to

$$M \ \frac{dy}{dt} + A(t)y = Bu \ , \ t \in [0,T], \tag{4.5.6b}$$

$$My(0) = Gu. \tag{4.5.6c}$$

Theorem 4.5.1. *The semi-discrete optimization problem (4.5.6a), (4.5.6b) admits a unique solution. If $y \in C^1([0,T], \mathbb{R}^N)$, $u \in U_{ad}$ is the optimal solution, there exists $p_Q \in C^1([0,T], \mathbb{R}^{N_\Omega})$ such that*

$$M_\Omega^T \ \frac{dp_Q}{dt} - A_\Omega(t)^T p_Q = -C_\Omega y_Q + \sum_{k=1}^2 (D_\Omega^{(k)})^T d_k \ , \ t \in [0,T], \tag{4.5.7a}$$

$$M_\Omega^T p_Q(T) = 0, \tag{4.5.7b}$$

holds true and the variational inequality

$$\Big(- G_\Omega^T p_Q(0) - \int_0^T B_\Omega(t)^T p_Q \ dt \Big) \cdot (v - u) \ge 0, \quad v \in U_{ad}, \tag{4.5.7c}$$

is satisfied, where $B_\Omega(t) = (0 B_{M,\Omega}(t))$, $0 \in \mathbb{R}^{(M-1) \times N_\Omega}$.

Proof. We introduce multipliers

$$p_Q \in C^1([0,T], \mathbb{R}^{N_\Omega}), \quad p_{\Sigma_\nu} \in C^1([0,T], \mathbb{R}^{N_{\Gamma_\nu}}), \ 1 \le \nu \le 2, \quad q_0 \in \mathbb{R}^N,$$

and consider the Lagrangian

$$L(y,u,p,q_0) := J(y,u) + \int_0^T p \cdot (M \ \frac{dy}{dt} + Ay - Bu) \ dt + q_0 \cdot (My(0) - Gu).$$

Integration by parts reveals

$$L(y,u,p,q_0) = J(y,u) + \int\limits_0^T y \cdot (-M^T \frac{dp}{dt} + A^T p)\ dt + - \int\limits_0^T p \cdot Bu\ dt$$
$$+ y(T) \cdot M^T p(T) - y(0) \cdot M^T p(0) + y(0) \cdot M^T q_0 - q_0 \cdot Gu.$$

The optimality condition $L_y(y,u,p,q_0) = 0$ gives $q_0 = p_Q(0)$ and shows that p_Q satisfies (4.5.7a), (4.5.7b). Moreover, the optimality condition $L_{q_0}(y,u,p,q_0) = 0$ yields $My(0) = Gu$ and together with $L_p(y,u,p,q_0) = 0$ implies that y satisfies (4.5.6b)-(4.5.6c). Finally, $(L_u(y,u,p,q_0)) \cdot (v-u) \geq 0$, $v \in U_{ad}$ results in (4.5.7c). □

4.6. Fully discrete optimal control problem

For the discretization in time of the semi-discrete optimal control problem (4.5.6a)-(4.5.6c) we consider a partition

$$0 =: t_0 < t_1 < \cdots < t_R := T, \quad R \in \mathbb{N},$$

of the time interval $[0,T]$ with step lengths $\tau_r := t_r - t_{r-1}$, $1 \leq r \leq R$. We approximate the ODE (4.5.6b) by the backward Euler scheme, split the integral in the objective functional (4.5.6a) into the sum over the subintervals (t_{r-1}, t_r) and use the quadrature formula $\int_{t_{r-1}}^{t_r} v dt \approx \tau_r v(t_r)$. Denoting by $y^r = (y_Q^r, y_{\Sigma_1}^r, y_{\Sigma_2}^r)^T$ approximations of $y = (y_Q, y_{\Sigma_1}, y_{\Sigma_2})^T$ at t_r, $0 \leq r \leq R$, and setting $y := (y^0, \cdots, y^R)^T$, $y_Q := (y_Q^0, \cdots, y_Q^R)^T$, $y_{\Sigma_\nu} := (y_{\Sigma_\nu}^0, \cdots, y_{\Sigma_\nu}^R)^T$, $1 \leq \nu \leq 2$, the fully discrete optimal control problem reads: Find $(y,u) \in \mathbb{R}^{(R+1)N} \times U_{ad}$ such that

$$\inf_{y,u} J(y,u) := \frac{1}{2} \sum_{r=1}^{R} \tau_r \Big((y_Q^r)^T C_\Omega y_Q^r - 2 \sum_{k=1}^{2} d_k^T D_\Omega^{(k)} y_Q^r + \sum_{k=1}^{2} d_k^T M_\Omega d_k \Big), \tag{4.6.1a}$$

subject to

$$My^r + \tau_r A(t_r) y^r = \tau_r Bu + My^{r-1},\ 1 \leq r \leq R, \tag{4.6.1b}$$
$$My^0 = Gu. \tag{4.6.1c}$$

Theorem 4.6.1. *The fully discrete optimization problem (4.6.1a)-(4.6.1c) admits a unique solution. If $y \in \mathbb{R}^{(R+1)N}$, $u \in U_{ad}$ is the optimal solution, there exists $p_Q = (p_Q^0, \cdots, p_Q^R)^T \in \mathbb{R}^{(R+1)N_\Omega}$ such that*

$$M_\Omega p_Q^{r-1} + \tau_r A_\Omega(t_{r-1})^T p_Q^{r-1} = M_\Omega p_Q^r + \tau_r (C_\Omega y_Q^r + \sum_{k=1}^{2} (D_\Omega^{(k)})^T d_k), \tag{4.6.2a}$$
$$M_\Omega p_Q^R = 0, \tag{4.6.2b}$$

and

$$\Big(- G_\Omega^T p_Q^0 - \sum_{r=0}^{R-1} \tau_{r+1} B_\Omega(t_r)^T p_Q^r \Big) \cdot (v - u) \geq 0 \quad , \quad v \in U_{ad}. \tag{4.6.2c}$$

Proof. The proof is the discrete analogue of Theorem 4.5.1. □

4.7. The algorithmic approach

We denote by $S : U_{ad} \to \mathbb{R}^{(R+1)N}$ the control-to-state map which assigns to an admissible control $u \in U_{ad}$ the solution $y \in \mathbb{R}^{(R+1)N}$ of the discrete state equation (4.6.1a), (4.6.1b). Then, the so-called control-reduced form of the fully discrete optimal control problem (4.6.1a)-(4.6.1c) reads

$$\inf_{u \in U_{ad}} J_{red}(u), \quad J_{red}(u) := J(S(u), u). \tag{4.7.1}$$

It follows from Theorem 4.6.1 that the gradient of the control-reduced objective functional is given by

$$\nabla J_{red}(u) = -G_\Omega^T p_Q^0 - \sum_{r=0}^{R-1} \tau_{r+1} B_\Omega(t_r)^T p^r. \tag{4.7.2}$$

Denoting by $P_{U_{ad}} : \mathbb{R}^M \to U_{ad}$ the pointwise projection onto the admissible control set U_{ad}, i.e.,

$$P_{U_{ad}}(w) = \begin{cases} u_{i,min} & \text{if } w_i \leq u_{i,min} \\ w_i & \text{if } u_{i,min} \leq w_i \leq u_{i,max} \\ u_{i,max} & \text{if } w_i \geq u_{i,max} \end{cases}, \quad 1 \leq i \leq M, \tag{4.7.3}$$

and given an initial control $u^{(0)} \in U_{ad}$, we solve (4.6.1a)-(4.6.1c) by the projected BFGS method with Armijo line search (cf., e.g., [Kel99, NW99])

$$u^{(\ell+1)} = P_{U_{ad}} \left(u^{(\ell)} - \alpha_\ell \, d_\ell \right), \ \ell \geq 0. \tag{4.7.4}$$

The search direction d_ℓ is defined as

$$d_\ell = -H_\ell^{-1} \nabla J_{red}(u^{(\ell)}), \ \ell \geq 0$$

where the approximate Hessian H_ℓ has been updated by the BFGS method:

$$\begin{aligned} H_\ell &= H_{\ell-1} + \frac{z_\ell z_\ell^T}{z_\ell^T s_\ell} - \frac{(H_{\ell-1} s_\ell)(H_{\ell-1} s_\ell)^T}{s_\ell^T H_{\ell-1} s_\ell}, \ \ell > 0 \\ z_\ell &= \nabla J_{red}(u^{(\ell)}) - \nabla J_{red}(u^{(\ell-1)}), \ \ell > 0 \\ s_\ell &= u^{(\ell)} - u^{(\ell-1)}, \ \ell > 0 \end{aligned}$$

and $H_0 = I$. The step length α_ℓ is determined by Armijo line search. It is chosen such that it satisfies the Armijo rule

$$J_{red}\left(P_{U_{ad}}(u^{(\ell)} - \alpha_\ell\, d_\ell)\right) \leq J_{red}(u^{(\ell)}) + c_1 \alpha_\ell d_k^T \nabla J_{red}(u^{(\ell)}) \tag{4.7.5}$$

where $0 < c_1 \ll 1$. This leads to the following algorithm:

Algorithm 3 Projected BFGS Method with Armijo Line Search

Input: TOL such that $TOL > 0$;
$u^{(0)} \in \mathbb{R}^M$ such that $u^{(0)} \in U_{ad}$

1: set $\ell := 0$
2: **repeat**
3: compute $y^{(\ell+1)} = S(u^{(\ell)}) \in \mathbb{R}^{(R+1)N}$ and $p_Q^{(\ell+1)} \in \mathbb{R}^{(R+1)N_\Omega}$ as the solutions of (4.6.1b), (4.6.1c) and (4.6.2a), (4.6.2b) ▷ with u in (4.6.1c) replaced with $u^{(\ell)}$
4: update the control by computing $u^{(\ell+1)}$ according to (4.7.4) with step length α_ℓ chosen by means of (4.7.5)
5: set $\ell = \ell + 1$
6: **until** $\|\nabla J_{red}(u^{(\ell-1)})\| < TOL$

Output: $(y^{(\ell)}, u^{(\ell)}, p^{(\ell)})$

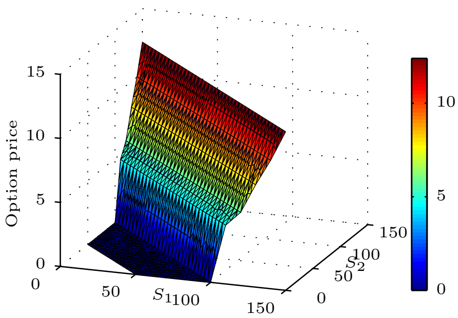
4.8. Numerical results

We provide documentation of computational results for the optimal control of European double barrier basket call options based on the numerical solution of the optimal control problem as outlined in section 4.5 and section 4.6. We have considered the case of $M = 5$ controls, i.e., $u = (u_1, \cdots, u_5)^T$ and various values of $d = (d_1, d_2)^T$ in the tracking-type objective functional of the optimal control problem. In the first example, we have used constant volatilities σ_1, σ_2 and interest rate r, whereas example 2 deals with the case of variable data. Finally, in order to illustrate the benefits of the optimal control approach, example 3 deals with a non-optimized European double barrier basket call option featuring a single cash-settlement at the upper barrier.

Example 1. In our first example we have chosen $d = (0.2, 0.2)^T$ in the objective functional. The complete data are given in Table 4.8.1 below.
For discretization in space, we have chosen a simplicial triangulation of Ω with $h = 5.0$ for both the state and the adjoint state and for discretization in time we have used a uniform time step of $\Delta t = 0.01$. The projected BFGS method with Armijo line search has been initialized with the initial control $u_0 = (26, 18, 5, 8, 10)^T$ and has been stopped when the norm of the gradient of the objective function became smaller than $TOL := 1.0E - 06$. The iteration terminated after 30 iterations with

Parameter	Notation	Value
d	Desired Delta	$(0.2, 0.2)^T$
M	Number of controls	5
K_{min}	Lower Barrier	50
K_{max}	Upper Barrier	150
K	Strike	100
T	Maturity	1
r	Interest Rate	0.05
σ_1	Volatility of asset 1	0.35
σ_2	Volatility of asset 2	0.20
ρ	Correlation between the assets	-0.5
$u_{i,min}$	Lower bound on the controls	0.0
$u_{i,max}$	Upper bound on the controls	50.0

Table 4.8.1.: Example 1: Data of the optimal control problem



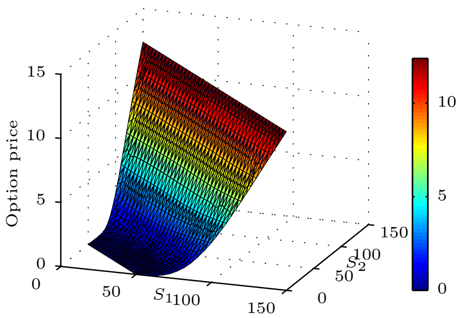


Figure 4.8.1.: Example 1: The option price, i.e. $y(\cdot, t)$, at maturity $t = 0$ (left) and at time to maturity $t = 0.5$ (right) for $u = u^*$.

the optimal control $u^* = (4.59, 5.71, 8.00, 10.03, 12.39)^T$. The corresponding state y^* is plotted for two different time instances in Figure 4.8.1.

The convergence history of the projected BFGS algorithm with Armijo line search is plotted in Figure 4.8.2. Here, ℓ stands for the iteration number, $J_{red}(u^{(\ell)})$ is the corresponding value of the objective functional, and $\|\nabla J_{red}(u^{(\ell)})\|$ refers to the norm of the gradient.

Example 2. The second example deals with the case of space-varying volatilities σ_1, σ_2, and time-varying interest rate r. The complete data are given in table 4.8.2 below.

We have used the same discretizations as in the first example. The initial control has been $u_0 = (50, 0, 50, 0, 50)^T$. As tolerance for the termination criterion we have used $TOL = 10^{-6}$. The computed optimal control is $u^* = (17.50, 18.98, 25.07, 29.40, 35.77)^T$.

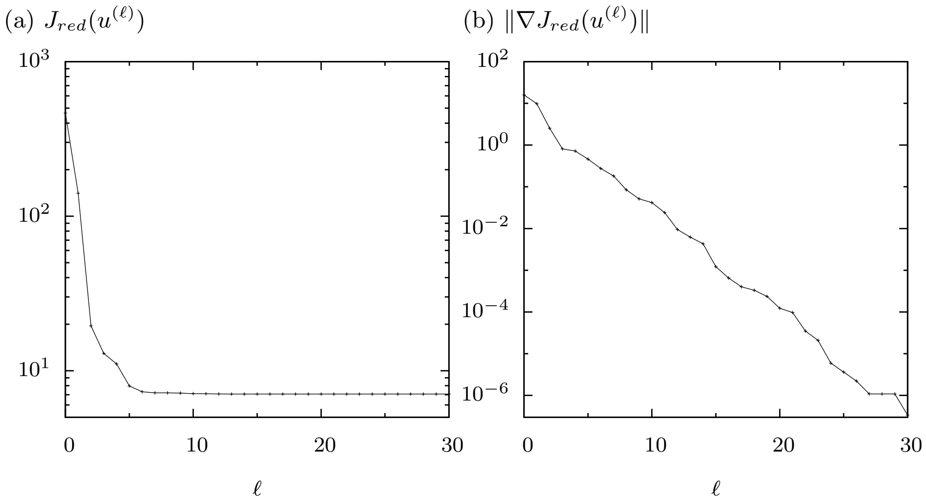


Figure 4.8.2.: Example 1: Convergence history: ℓ the number of projected BFGS iteration.

The corresponding state at maturity and at time to maturity $t = 0.5$ is plotted in Figure 4.8.3.

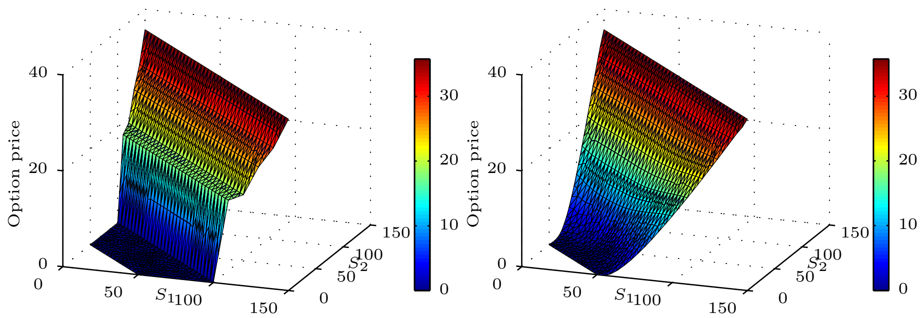


Figure 4.8.3.: Example 2: The option price, i.e. $y(\cdot, t)$, at maturity $t = 0$ (left) and at time to maturity $t = 0.5$ (right) for $u = u^*$.

In Figure 4.8.4 the convergence history of the projected BFGS algorithm with Armijo line search is plotted. The value of the objective functional at optimality is significantly higher than in the first example due to the higher values of d_1 and d_2. On the other hand, the projected BFGS method with Armijo line search performed similarly. Indeed, the termination criterion was reached after 46 iterations.

Example 3. In order to illustrate the benefits of optimized versus non-optimized European double barrier basket call options, we present the numerical results for a non-optimized call option with a single cash settlement of 10 at the upper barrier,

Parameter	Notation	Value
d	Desired Delta	$(0.5, 0.5)^T$
M	Number of controls	5
K_{min}	Lower Barrier	50
K_{max}	Upper Barrier	150
K	Strike	100
T	Maturity	1
$r(t)$	Interest Rate	$0.02 \cdot t + 0.08 \cdot (1 - t)$
$\sigma_1(S, t)$	Volatility of asset 1	$0.75 \cdot (1 - (S_1 + S_2$ $-100)/50)^2$
$\sigma_2(S, t)$	Volatility of asset 2	$0.75/2 \cdot (1 - (S_1 + S_2$ $-100)/50)^2$
ρ	Correlation between assets	-0.5
$u_{i,min}$	Lower bound on the controls	0.0
$u_{i,max}$	Upper bound on the controls	50.0

Table 4.8.2.: Example 2: Data of the optimal control problem

but otherwise the same data as in the previous example.
In case the option is still in the money at maturity, its price corresponds to that of a plain vanilla European call option, as can be seen in Figure 4.8.5 (left). Otherwise, there are significant differences as displayed in Figure 4.8.5 (right). In fact, in comparison with European double barrier basket call options featuring optimized cash settlements and aiming at a constant delta, the delta is varying considerably and can even take negative values.

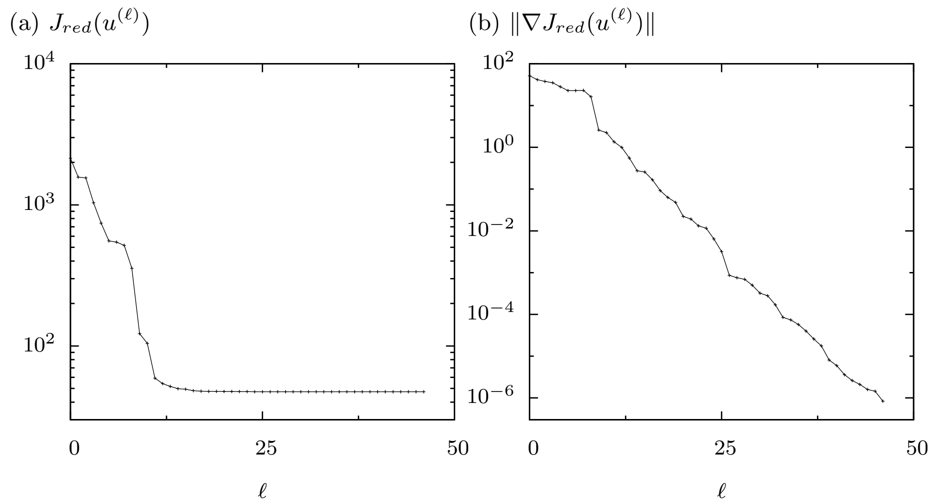


Figure 4.8.4.: Example 2: Convergence history: ℓ the number of the projected BFGS iteration.

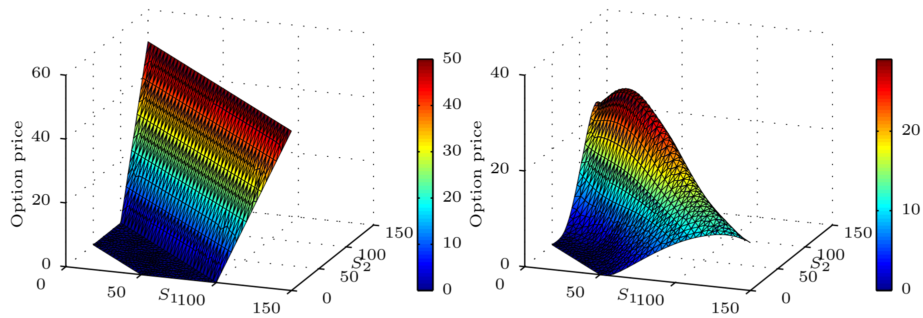


Figure 4.8.5.: Example 3: Option price at maturity (left) and time to maturity $t = 0.5$ (right).

Concluding remarks

The dissertation has contributed numerical methods for optimization in finance.

- We have proposed a Monte Carlo method. The method allows to approximate hedging strategies which are optimal in the framework of local risk minimization in discrete time.
- We have proposed an optimal control approach for the optimization of European double barrier options on a basket of two assets. The approach allows to control the payoff and the rebate at the upper barrier such that the delta of the option is as close as possible to a predefined constant.

Optimal dynamic hedging

Chapters 1-3 have been on optimal dynamic hedging with focus on the hedged Monte Carlo method. Chapter 1 has been of theoretical nature. Chapters 2 and 3 have been on the application of the hedged Monte Carlo method.

In chapter 1 the hedge problem has been formulated in the framework of local risk minimization. The setting has been fairly general: N hedging instruments, minimal requirements on the stochastic process X which models the price of the hedging instrument and minimal requirements on the random variable H which models the hedging objective. The hedged Monte Carlo method has been described subsequently without further restrictions.
The formulation of the problem could be generalized further: one may allow payment streams for instance. In this case H is replaced by a stochastic process $(H_k)_{k=0}^K$. This has been done recently by Schweizer [Sch08] for local risk minimization in continuous time. When it is done for local risk minimization in discrete time one may address the discretization of the problem/the generalization of the hedged Monte Carlo method. The paper of Coleman et al. [CLL07] may contain useful information. They studied the discrete hedging of American-type options using local risk minimization. The work, however, is limited to the case where there is only one hedging instrument and where this is the underlying asset of the hedging objective. Furthermore they considered a binomial tree model.
The risk process $R = (R_k)_{k=0}^K$ has been defined as the expected quadratic cost increments (quadratic local risk minimization). Other definitions are possible. In [CLL07] for instance quadratic local risk minimization is compared with (constrained) piecewise linear local risk minimization. In [PB03] the risk process is a function of the cost increment.

Chapters 2 and 3 have been devoted to practical aspects of the hedged Monte Carlo method/algorithm. The problem formulation requires the specification of two processes: the process $F = (F_k)_{k=0}^K$ which generates the filtration and the process $X = (X_k)_{k=0}^K$ which models the price of the hedging instruments. The processes may coincide but they do not in general, cf. chapter 2 and 3. As a practical matter one should choose the processes such that one is capable to simulate, i.e. to compute, many realizations, I, of the processes in little time.
In a subsequent step various parameters have to be chosen for the discretization of the problem. The determination of an appropriate basis is most involved since one has to fix the basis size and the type of the basis functions. In general, increasing the basis size does not improve the approximation quality. Once the basis has been chosen, setting I is easy. In general, it is subject to the required precision and the available computing time.
B-splines of order 3 respectively of order $(3, 2)$ turned out to be a good choice. We believe that this does not only hold true for the specific problems investigated.
The discretization problem of choosing the basis appropriately can be viewed from the point of least squares regression. The task is to regress certain curves/surfaces. The more complicated the shape of the curves/surfaces the more delicate it is to choose the basis functions. This can be seen particularly well in figure 2.3.1. Basically, there are two approaches to choose a basis: either one looks for a very flexible basis or one looks for a very problem-adapted basis. We have made the decision for the first approach. B-splines are not all adapted to the hedging problem. Hence, they are supposed to be robust with respect to model parameters variations. A drawback is that the B-spline bases are relatively large. This is particular true if the basis functions are functions of more than two variables.
We tried to choose a basis adapted to the problem as well. Doing this by hand has not led to success. A strategy/method to accomplish this task would be required.

From an economic point of view and what concerns the risk sources, local risk minimization has a more global character compared to Black-Scholes greek hedging. We observed this in chapter 3. With the strategy obtained by the hedged Monte Carlo algorithm both price and volatility risk are hedged in part. To which extend this happened is not known but it is optimal in the sense of the local risk minimization. In the case of Black-Scholes greek hedging, each hedging instrument is used in general to hedge a specific greek. If there are two hedging instruments, this can be for instance delta and gamma or delta and vega.

For our second model problem, chapter 3, we have observed that the efficiency of the hedged Monte Carlo method depends on the number of hedging instruments or more precisely on the residual risk. The less risk the less variance. This led to speed up of factor 7 in the double hedged case.

The hedging objective has been a vanilla put option in chapter 2 and 3. The framework, however, is general enough that one could have other hedging objectives as long as they have European exercise style, e.g. Asian options, lookback options, basket options etc. If the exercise style is not European, e.g. American, Bermudan, Canary, etc., the framework and the hedged Monte Carlo method have to be generalized. Practical guidance may provide the paper of Potters, Bouchaud and Sestovic [PBS01]. They provided numerical results on the valuation of American options.

For further variance reduction one may incorporate the method of antithetic variates or possibly other variance reduction techniques, see [BBG97, Gla03].

As a final remark, one can use historical data instead of generating realizations of the processes which model the price of the hedging instruments. We refer to [PBS01] for some numerical results in the pricing context.

Optimal control of European double barrier basket options

Chapter 4 has been about optimal control of European double barrier basket options.

The problem has been formulated in an appropriate function space setting based on the weak formulation of the Black-Scholes equation. The discretization in the spatial domain by conforming P1 finite elements led to a semi-discrete formulation of the optimal control problem. Its fully discrete form has been obtained after discretization of the temporal domain by the implicit Euler scheme. We showed for each of the three formulations that the optimal control problem has an unique solution and that there exists an adjoint state which satisfies an associated adjoint problem and an associated variational inequality. We proposed an algorithmic approach to the solution of the fully discrete optimal control problem based on the projected BFGS method with Armijo line search. Numerical results have been provided for constant and non-constant parameters (the interest rate respectively the volatilities of the underlying assets) demonstrating the successful application of the algorithmic approach and the benefit of optimized double barrier basket options with respect to non-optimized double barrier basket options.

The solution of the fully discrete optimal control problem is computationally challenging if the coefficients are time dependent or if the discretization parameters (the spatial and the temporal mesh size) are chosen small. In the last case, large-scale systems of differential resp. difference equations have to be solved. In order to reduce the computing time one may consider parallelism or model order reduction.

Parallel computing is based on task decomposition. This can be achieved for instance by decomposing the spatial (see e.g. [TW04, QV99]) and/or temporal domain. For parallelism across the time one may use the parareal algorithm [LMT01, MT02].

Model reduction seeks to replace an original complex model by a simpler one. The complex model can be a large-scale system of equations and the problem is then to set up a system of smaller scale which has approximately the same response characteristics [Ant05]. Two model reduction techniques which may be employed in the context of optimal control of double barrier options are proper orthogonal decomposition and balanced truncation.

Balanced truncation preserves asymptotic stability and there are error bounds on the discrepancy between the outputs of the full and the reduced order system. The applicability, however, is limited to time invariant systems. This means that the coefficients (the interest rate and the volatilities) have to be constant with respect to time. We refer to [AHHS10] for the application of balanced truncation in the context of optimal control and to [HRA11] for the application of balanced truncation to systems with inhomogeneous initial condition.
For proper orthogonal decomposition stability and error bounds have not been proven yet. In contrast to balanced truncation, time invariance of the system is not required [Vol]. In the financial context proper orthogonal decomposition has been used for instance by Sachs and Schu [SS08].

Another approach to reduce the computational load could be using the reduced basis method, see e.g. Maday et al. [MPT02] or for an application in the financial context see [Pir09, CLP10].

The final remark concerns the optimization of other types of options. Despite this may involve a series of modifications, the basic optimal control approach will hold. We considered a basket of two assets for computational and notational convenience. If the basket consists of three assets, one has to solve the Black-Scholes equations on three-, two- and one-dimensional domains.

Appendix

A. Optimal dynamic hedging

A.1. Basis functions

A.1.1. Laguerre polynomials

Laguerre polynomials can be defined by repeated differentiation.

Definition A.1.1. The k-th Laguerre polynomial $L_k : \mathbb{R}_+ \to \mathbb{R}$, $k \in \mathbb{N}_0$, is defined by

$$L_k(x) := \frac{e^x}{k!}\frac{d^k}{dx^k}\left(x^k e^{-x}\right) \qquad x > 0. \tag{A.1.1}$$

The following theorem is useful for implementation.

Proposition A.1.2. *For $x \in \mathbb{R}_+$ there holds*

$$\begin{aligned} L_0(x) &= 1, \\ L_1(x) &= 1 - x, \\ L_{k+1}(x) &= \frac{2k+1-x}{k+1}L_k(x) - \frac{k}{k+1}L_{k-1}(x), \quad k > 0. \end{aligned}$$

A proof of the proposition can be found in [AS64].

Proposition A.1.3. *Laguerre polynomials are orthogonal with respect to*

$$< f, g >_w := \int_0^\infty f(x)g(x)e^{-x}dx, \qquad f, g : \mathbb{R}_+ \to \mathbb{R}$$

i.e.

$$< L_i, L_j >_w = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} \quad i, j \in \mathbb{N}_0.$$

We refer to [AS64] for a proof of the proposition.

Definition A.1.4. The k-th weighted Laguerre polynomial $L_k^w : \mathbb{R}_+ \to \mathbb{R}$, $k \in \mathbb{N}_0$, is defined by

$$L_k^w(x) := e^{-x/2} \cdot L_k(x) \quad x > 0.$$

Lemma A.1.5. *The weighted Laguerre polynomials are orthogonal with respect to*

$$< f, g > := \int_0^\infty f(x)g(x)dx, \qquad f, g : \mathbb{R}_+ \to \mathbb{R}.$$

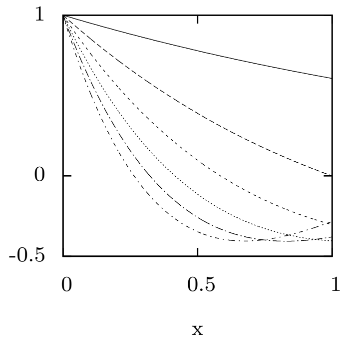


Figure A.1.1.: Plot of L_k^w for $k = 0, \dots, 5$

Figure A.1.1 shows a plot of the first six weighted Laguerre polynomials.

Definition A.1.6. The k-th Laguerre polynomial $L_{k_1,k_2} : \mathbb{R}_+^2 \to \mathbb{R}$, $k \in \mathbb{N}_0$, is defined by

$$L_{k_1,k_2}(x_1, x_2) := L_{k_1}(x_1)L_{k_2}(x_2) \quad x_1, x_2 > 0$$

and the k-th weighted Laguerre polynomial $L_{k_1,k_2}^w : \mathbb{R}_+^2 \to \mathbb{R}$, $k \in \mathbb{N}_0$, is defined by

$$L_{k_1,k_2}^w(x_1, x_2) := e^{-(x_1+x_2)/2} \cdot L_{k_1,k_2}(x_1, x_2) \quad x_1, x_2 > 0.$$

A.1.2. Basis splines

Let $a, b \in \mathbb{R}$ with $a < b$ and $n \in \mathbb{N}$, $n > 1$.

Definition A.1.7. A partition of $[a, b]$ is a set $\lambda = \{\lambda_i\}_{i=0}^{n+1}$ with $\lambda_i \in \mathbb{R}$, $\lambda_i < \lambda_{i+1}$, $\lambda_0 = a$ and $\lambda_{n+1} = b$.

Let λ be a partition of $[a, b]$ and $k \in \mathbb{N}_0$.

Definition A.1.8. i. A function $f \in C^{k-2}[a, b]$ is called spline of order k if f is a polynomial of order k on $[\lambda_i, \lambda_{i+1}]$, $i = 0, \dots, n$.
ii. The space of all splines of order k is denoted by $S_{k,\lambda}$.

Proposition A.1.9. *The dimension of $S_{k,\lambda}$ is $n + k$.*

The proposition is proven in [DH08].
Basis splines or short B-splines form a special set of splines. This set constitutes a basis of $S_{k,\lambda}$. They are of numerical interest since they have minimal support and can be evaluated in a stable way.

Definition A.1.10. Let $d = n + k$, $k < d$. Let τ_i, $i = 1, \dots, d + k$, be such that

$$\tau_i = a \qquad \text{for } i = 1, \dots, k, \tag{A.1.2a}$$
$$\tau_i = \lambda_{i-k} \qquad \text{for } i = k + 1, \dots, d, \tag{A.1.2b}$$
$$\tau_i = b \qquad \text{for } i = d + 1, \dots, d + k. \tag{A.1.2c}$$

Let

$$N_{i,1}(x) := \begin{cases} 1, & x \in [\tau_i, \tau_{i+1}] \\ 0, & x \in [a,b] \setminus [\tau_i, \tau_{i+1}] \end{cases} \tag{A.1.3}$$

for $i = 1, \dots, d+k-1$ and let

$$N_{i,j}(x) := \frac{x - \tau_i}{\tau_{i+j-1} - \tau_i} N_{i,j-1}(x) + \frac{\tau_{i+j} - x}{\tau_{i+j} - \tau_{i+1}} N_{i+1,j-1}(x). \tag{A.1.4}$$

for $j = 2, \dots, k$, $i = 1, \dots, d+k-j$. Then, $N_{i,k}$, $1 \le i \le d$, is called the i-th B-spline of order k.

Proposition A.1.11.

$$S_{k,\lambda} = span\{N_{1,k}, \dots, N_{d,k}\}.$$

We refer for a proof of the proposition to [DH08].

Figure A.1.2 provides visualization of B-splines of order 1-4.

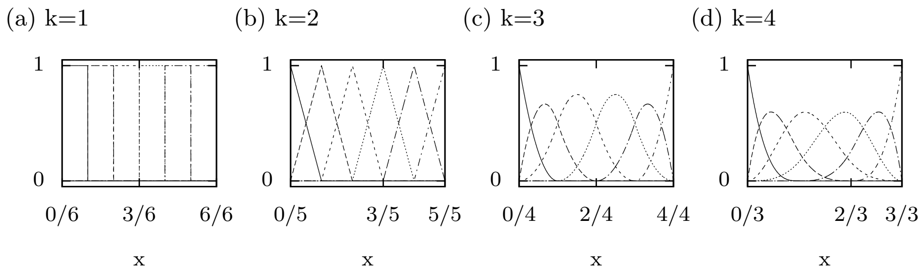


Figure A.1.2.: Plots of $N_{i,k}$ for $i = 1, \dots, 6$ with respect to $\lambda = \{j/(6+1-k)\}_{j=0}^{6+1-k}$.

Definition A.1.12. For $j = 1, 2$, let $a_j, b_j \in \mathbb{R}$ with $a_j < b_j$, let λ^j be a partition of $[a_j, b_j]$, let $k_j, d_j \in \mathbb{N}$ with $k_j < d_j$ and let $N^j_{1,k_j}, \dots, N^j_{d_j,k_j}$ be B-splines with

$$\text{span}\{N^j_{1,k_j}, \dots, N^j_{d_j,k_j}\} = S_{k_j,\lambda^j}.$$

Let $N_{i_1,i_2,k_1,k_2} : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}$, $i_j = 1, \dots, k_j$, $j = 1, 2$, be defined by

$$N_{i_1,i_2,k_1,k_2}(x_1, x_2) = N^1_{i_1,k_1}(x_1) N^2_{i_1,k_2}(x_2) \quad (x_1, x_2) \in [a_1, b_1] \times [a_2, b_2].$$

The functions N_{i_1,i_2,k_1,k_2}, $i_j = 1, \dots, k_j$, $j = 1, 2$ are called surface splines or two-dimensional B-splines of order (k_1, k_2).

A.2. Estimator

Let $(\Omega, \mathcal{F}, P)$ be a probability space and $X : \Omega \to \mathbb{R}$ be a random variable with $E[X^4] < \infty$. Let $X_1, \ldots, X_L$, $L \in \mathbb{N}$, be L realizations of X, i.e. i.i.d. random variables with $E[X_1^4] < \infty$.

To estimate sample mean, standard deviation, skewness and kurtosis we define the following estimators.

Definition A.2.1.

$$\bar{X} := \mathrm{Mean}(X) := \frac{1}{L} \sum_{\ell=1}^{L} X_\ell \tag{A.2.1a}$$

and with

$$\mu_n(X) := \frac{1}{L} \sum_{\ell=1}^{L} (X_\ell - \bar{X})^n, \quad n = 2, 3, 4$$

let

$$\mathrm{StDev}(X) := (\mu_2(X))^{1/2} \tag{A.2.1b}$$

$$\mathrm{Skew}(X) := \frac{\mu_3(X)}{(\mu_2(X))^{3/2}} \tag{A.2.1c}$$

$$\mathrm{Kurt}(X) := \frac{\mu_4(X)}{(\mu_2(X))^2} - 3. \tag{A.2.1d}$$

A shorter notation for Mean is m and for StDev it is s.

A.3. Black-Scholes formula

Let $\varphi : \mathbb{R} \cup \{\pm\infty\} \to \mathbb{R}$ and $N : \mathbb{R} \cup \{\pm\infty\} \to \mathbb{R}$ be functions defined by

$$\varphi(x) := \begin{cases} \frac{1}{2\pi} e^{-\frac{x^2}{2}} & \text{if } x \in \mathbb{R} \\ 0 & \text{if } x = \pm\infty \end{cases}$$

and

$$N(x) := \begin{cases} 0 & \text{if } x = -\infty \\ 1 & \text{if } x = +\infty \\ \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-\frac{x^2}{2}} dx & \text{if } x \in \mathbb{R}. \end{cases}$$

Definition A.3.1. For $S, \tau, K, \sigma > 0$ and $r \geq 0$ let

$$\texttt{bs\_put\_price}(S, \tau; K, r, \sigma) := Ke^{-r\tau} N(-d_2) - SN(-d_1), \tag{A.3.1}$$

$$\texttt{bs\_put\_delta}(S, \tau; K, r, \sigma) := -N(-d_1), \tag{A.3.2}$$

$$\texttt{bs\_gamma}(S, \tau; K, r, \sigma) := \frac{\varphi(d_1)}{S\sigma\sqrt{\tau}}, \tag{A.3.3}$$

with

$$d_1 := \frac{\ln(\frac{S}{K}) + (r + \frac{\sigma^2}{2})\tau}{\sigma\sqrt{\tau}}, \qquad d_2 := d_1 - \sigma\sqrt{\tau}.$$

The definitions are explained in many textbooks, we refer for instance to [Hul08], [Wil07] or [WHD95].

B. Mixed stochastic/deterministic methods for option pricing

We have investigated the combination of Monte Carlo, quadrature and PDE methods for pricing basket options in the Black-Scholes model. The derived methods are introduced for European put options on a basket of three assets. Numerical results are provided.

The work has been realized in cooperation with Olivier Pironneau and Denis Talay.

Contents

B.1. Introduction

Since the pioneering work of Phelim Boyle [Boy77] Monte Carlo methods have entered and shaped mathematical finance like barely any other method. They are often appreciated due to their flexibility and applicability in high dimensions although they go hand in hand with a number of drawbacks: error terms are probabilistic and a high level of accuracy can be computationally burdensome to achieve. In low dimensions deterministic methods as quadrature and quadrature based methods are strong competitors: They often allow deterministic error estimation and give precise results.

We propose several methods for pricing basket options in a Black-Scholes framework. The methods are based on a combination of Monte Carlo, quadrature and PDE methods. The key idea is to uncouple the underlying system of stochastic differential equations (SDEs) and then to apply the last-mentioned methods appropriately.

The problem is formulated in section B.2. The system of SDEs which describes the price of the underlying assets is uncoupled in section B.3. We outline various combinations of Monte Carlo, quadrature and PDE methods in section B.4. Numerical results are presented in section B.5. We have used mixed methods for pricing European put options on a basket of three respectively five assets. In particular we have used finite elements for the discretization of the Black-Scholes PDE in the spatial domain. In section B.6 we draw some conclusions.

B.2. Problem formulation

We consider an option, P, on three assets whose dynamics are determined by the following system of stochastic differential equations: for $i = 1, 2, 3$

$$dS_{i,t} = S_{i,t}(rdt + dW_{i,t}), \quad t > 0 \tag{B.2.1}$$

with initial condition $S_{i,t=0} = S_{i,0}$, $S_{i,0} \in \mathbb{R}_+$. The parameter r, $r \in \mathbb{R}_{\geq 0}$, is constant and $W_i := \sum_{j=1}^{3} a_{ij} B_j$ are linear combinations of standard Brownian motions B_j such that

$$\mathrm{Cov}\left[W_{i,t}, W_{j,t}\right] = \rho_{ij}\sigma_i\sigma_j t, \quad t > 0.$$

We further assume that $\Xi := (\rho_{ij}\sigma_i\sigma_j)_{i,j=1}^{3}$ is symmetric positive definite with

$$\rho_{ii} = 1 \quad \text{and} \quad \rho_{ij} \in (-1, 1) \text{ otherwise.}$$

The coefficients a_{ij}, $a_{ij} \in \mathbb{R}$, have to be chosen such that

$$\begin{aligned}
\mathrm{Cov}\left[W_{i,t}, W_{j,t}\right] &= E\left[W_{i,t}W_{j,t}\right] \\
&= E\left[(a_{i1}B_{1,t} + a_{i2}B_{2,t} + a_{i3}B_{3,t})(a_{j1}B_{1,t} + a_{j2}B_{2,t} + a_{j3}B_{3,t})\right] \\
&= a_{i1}a_{j1}E\left[B_{1,t}^2\right] + a_{i2}a_{j2}E\left[B_{2,t}^2\right] + a_{i3}a_{j3}E\left[B_{3,t}^2\right] \\
&= (a_{i1}a_{j1} + a_{i2}a_{j2} + a_{i3}a_{j3})t, \quad t > 0
\end{aligned}$$

or equivalently such that

$$AA^T = \Xi$$

where $A := (a_{ij})_{i,j=1}^3$. Without loss of generality we may set the strict upper triangular components of A to zero and find

$$A = \begin{pmatrix} \sigma_1 & 0 & 0 \\ \sigma_2\rho_{21} & \sigma_2\sqrt{1-\rho_{12}^2} & 0 \\ \sigma_3\rho_{31} & \sigma_3\frac{\rho_{32}-\rho_{21}\rho_{31}}{\sqrt{1-\rho_{12}^2}} & \sigma_3\sqrt{1-\rho_{31}^2-\left(\frac{\rho_{32}-\rho_{21}\rho_{31}}{\sqrt{1-\rho_{12}^2}}\right)^2} \end{pmatrix}.$$

The option P has maturity T, $T \in \mathbb{R}_+$, strike K, $K \in \mathbb{R}_+$ and payoff function $\varphi : \mathbb{R}_+^3 \to \mathbb{R}$,

$$\varphi(x) = \left(K - \sum_{i=1}^{3} x_i\right)^+ \qquad x = (x_1, x_2, x_3)^T \in \mathbb{R}_+^3.$$

The Black-Scholes price of P at time 0 is

$$P_0 = e^{-rT}E^*\left[\left(K - \sum_{i=1}^{3} S_{i,T}\right)^+\right] \tag{B.2.2}$$

where E^* denotes the expectation with respect to the risk-neutral measure.

B.3. The uncoupled system

In order to combine different types of methods (Monte Carlo, quadrature and/or PDE methods) we will uncouple the SDEs in (B.2.1). We start with a change of variable to logarithmic prices. Let $s_{i,t} := log(S_{i,t})$, $i = 1, 2, 3$, then Itô's lemma shows that

$$ds_{i,t} = r_i dt + dW_{i,t} \quad t > 0 \tag{B.3.1}$$

with initial condition $s_{i,t=0} = s_{i,0} := log(S_{i,0})$. The parameters r_i, $i = 1, 2, 3$, have been defined as $r_i = r - \frac{a_{i1}^2}{2} - \frac{a_{i2}^2}{2} - \frac{a_{i3}^2}{2} = r - \frac{\sigma_i^2}{2}$. In the rest of this section the time index of any object is omitted to simplify the notation.

We note that equation (B.3.1) can be written as

$$\begin{pmatrix} ds_1 - r_1dt \\ ds_2 - r_2dt \\ ds_3 - r_3dt \end{pmatrix} = \begin{pmatrix} a_{11} & 0 & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} dB_1 \\ dB_2 \\ dB_3 \end{pmatrix}.$$

Then, uncoupling reduces to Gaussian elimination. Using the Frobenius matrices

$$F_1 := \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 \\ -\frac{a_{31}}{a_{11}} & 0 & 1 \end{pmatrix} \quad \text{and} \quad F_2 := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{a_{32}}{a_{22}} & 1 \end{pmatrix}$$

we write

$$F_2F_1(ds + rdt) = \mathrm{Diag}(a_{11}, a_{22}, a_{33})dB$$

where $s = (s_1, s_2, s_3)^T$, $r = (r_1, r_2, r_3)^T$ and $B = (B_1, B_2, B_3)^T$. We set $L^{-1} := F_2F_1$ and define

$$\tilde{s} := L^{-1}s \quad \text{and} \quad \tilde{S} := e^{L^{-1}s}.$$

Remark B.3.1. i) The processes $\tilde{s}_1$, $\tilde{s}_2$ and $\tilde{s}_3$ are independent of each other; analog $\tilde{S}_1$, $\tilde{S}_2$ and $\tilde{S}_3$.
ii) Let $\tilde{r} := L^{-1}r$ then

$$d\tilde{s} = \tilde{r}dt + \mathrm{Diag}(a_{11}, a_{22}, a_{33})dB.$$

iii) The coupled system expressed in terms of the uncoupled system is $s = L\tilde{s}$.
iv) In the next section we will make use of the triangular structure of $L = (L_{ij})_{i,j=1}^3$ and $L^{-1} = ((L^{-1})_{ij})_{i,j=1}^3$,

$$L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{a_{21}}{a_{11}} & 1 & 0 \\ \frac{a_{31}}{a_{11}} & \frac{a_{32}}{a_{22}} & 1 \end{pmatrix} \quad \text{and} \quad L^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -\frac{a_{21}}{a_{11}} & 1 & 0 \\ \frac{a_{21}a_{32}}{a_{11}a_{22}} - \frac{a_{31}}{a_{11}} & \frac{a_{32}}{a_{22}} & 1 \end{pmatrix}$$

v) The notation has been symbolic and the derivation heuristic.

B.4. Mixed methods

We describe nine combinations of Monte Carlo, quadrature and/or PDE methods.

Convention B.4.1. If Z is a stochastic process, we denote by Z^m, $m \in \mathbb{N}$, a realization of the process. Let N_{MC}, $N_{MC} \in \mathbb{N}$, stand for a fixed number of Monte Carlo drawings.

Basic methods

i) MC3 method: Simulate N_{MC} trajectories of (S_1, S_2, S_3). An approximation of the option price P_0 is

$$P_0^a := e^{-rT}\frac{1}{N_{MC}}\sum_{m=1}^{N_{MC}} \varphi(S_{1,T}^m, S_{2,T}^m, S_{3,T}^m).$$

ii) QUAD3 method: In order to use a quadrature formula we replace the risk neutral measure in

$$P_0 = e^{-rT} E^* \left[\left(K - e^{(L\tilde{s}_T)_1} - e^{(L\tilde{s}_T)_2} - e^{(L\tilde{s}_T)_3} \right)^+ \right]$$

by the Lebesgue-measure. Note,

$$\tilde{s}_{i,t} \sim N(\mu_{i,t}, a_{ii}^2 t), \quad 1 \le i \le 3$$

where $\mu_{i,t} = \tilde{s}_{i,0} + \tilde{r}_i t$. Let $f_{i,t}$ be the density of $\tilde{s}_{i,t}$, i.e.

$$f_{i,t}(x_i) = \frac{1}{\sqrt{2\pi} a_{ii} \sqrt{t}} e^{-\frac{1}{2}\left(\frac{x_i - \mu_{i,t}}{a_{ii}\sqrt{t}}\right)^2}, \quad x_i \in \mathbb{R}, \quad 1 \le i \le 3.$$

Due to the independence of $\tilde{s}_{1,t}$, $\tilde{s}_{2,t}$ and $\tilde{s}_{3,t}$, the density of

$$\left(K - e^{(L\tilde{s}_T)_1} - e^{(L\tilde{s}_T)_2} - e^{(L\tilde{s}_T)_3} \right)^+$$

is

$$(x_1, x_2, x_3) \mapsto f_{1,T}(x_1) f_{2,T}(x_2) f_{3,T}(x_3), \quad (x_1, x_2, x_3) \in \mathbb{R}^3.$$

The formula for the option price becomes

$$P_0 = e^{-rT} \int\limits_{\mathbb{R}^3} \left(K - e^{(Lx)_1} - e^{(Lx)_2} - e^{(Lx)_3} \right)^+ f_{1,T}(x_1) f_{2,T}(x_2) f_{3,T}(x_3) dx.$$

Now, a quadrature formula can be used to compute the integral.

The methods which are based on a combination of quadrature and some other method will be presented for the case where the trapezoidal rule is used. Next we show how the trapezoidal rule can be used to compute the integral. This allows us to introduce the notation for the description of methods which are combinations of quadrature and some other method(s).

To compute the integral we truncate the domain of integration to κ, $\kappa \in \mathbb{N}$, standard deviations around the means $\mu_{1,T}$, $\mu_{2,T}$ and $\mu_{3,T}$. Let

$$\begin{aligned} x_{i,0} &= \mu_{i,T} - \kappa a_{ii}^2 \\ x_{i,n} &= x_{i,0} + n\delta x_i, \quad n = 1, \ldots, N_Q, \end{aligned}$$

$1 \le i \le 3$, where $\delta x_i = \frac{2\kappa}{N_Q}$, $N_Q \in \mathbb{N}$.

The option price P_0 is then approximated by

$$P_0^a := e^{-rT} \sum_{n_1, n_2, n_3 = 1}^{N} \left(\prod_{i=1}^{3} \chi_{n_i} \delta x_i f_{i,T}(x_{i,n_i}) \right) \left(K - e^{(Lx_n)_1} - e^{(Lx_n)_2} - e^{(Lx_n)_3} \right)^+$$

where $x_n := (x_{1,n_1}, x_{2,n_2}, x_{3,n_3})^T$ and

$$\chi_n = \begin{cases} 0.5 & \text{if } n = 0 \text{ or } n = N_Q \\ 1 & \text{otherwise.} \end{cases}$$

Combination of two methods

iii) MC2-PDE1 method: Note,

$$\begin{aligned}P_0 &= e^{-rT} E^* \left[\left(K - S_{1,T} - S_{2,T} - S_{1,T}^{-2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}} \tilde{S}_{3,T} \right)^+ \right] \\ &= e^{-rT} E^* \left[\left(\bar{K} - \tilde{\tilde{S}}_{3,T} \right)^+ \right]\end{aligned}$$

where

$$\bar{K} := K - S_{1,T} - S_{2,T}$$

and $\tilde{\tilde{S}}_3$ is the solution of the stochastic initial value problem

$$\begin{aligned}d\tilde{\tilde{S}}_{3,t} &= \tilde{\tilde{S}}_{3,t} \left(\tilde{\tilde{r}}_3 dt + a_{33} dB_{3,t} \right) \\ \tilde{\tilde{S}}_{3,t=0} &= \alpha \tilde{S}_{3,0}\end{aligned}$$

with parameters $\tilde{\tilde{r}}_3 := \tilde{r}_3 + \frac{a_{33}^2}{2}$ and $\alpha = S_{1,T}^{-2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}}$.
The method is then: Simulate N_{MC} realizations of (S_1, S_2) and set $\bar{K}^m = K - S_{1,T}^m - S_{2,T}^m$ and $\alpha^m = {S_{1,T}^m}^{-2(L^{-1})_{31}} {S_{2,T}^m}^{-(L^{-1})_{32}}$. Compute an approximation of P_0 by

$$P_0^a := \frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} u(x_3, t; \bar{K}^m)_{|x_3 = \alpha^m \tilde{S}_{3,0}, t=T}$$

where u is the solution of the initial value problem for the one dimensional Black-Scholes PDE with parametrized (β) initial condition

$$\frac{\partial u}{\partial t} - \frac{(a_{33}x_3)^2}{2} \frac{\partial^2 u}{\partial x_3^2} - \tilde{\tilde{r}}_3 x_3 \frac{\partial u}{\partial x_3} + \tilde{\tilde{r}}_3 u = 0 \qquad \text{in } \Omega \times (0,T) \qquad \text{(B.4.1a)}$$

$$u(t=0) = u_0 \qquad \text{in } \Omega \qquad \text{(B.4.1b)}$$

where $\Omega = \mathbb{R}_+$ and

$$u_0(x_3; \beta) := (\beta - x_3)^+, \quad x_3 > 0.$$

iv) QUAD2-PDE1 method: Note,

$$P_0 = e^{-rT} \int_{\mathbb{R}^2} E^* \left[\left(K - e^{L_{11}x_1} - e^{L_{21}x_1 + L_{22}x_2} - e^{L_{31}x_1 + L_{32}x_2} e^{L_{33}\tilde{s}_{3,T}} \right)^+ \right]$$
$$f_{1,T}(x_1) f_{2,T}(x_2) dx_1 dx_2.$$

The option price P_0 is approximated by

$$P_0^a := \sum_{n_1,n_2=1}^{N_Q} \left(\prod_{i=1}^{2} \chi_{n_i} \delta x_i f_{i,T}(x_{i,n_i}) \right) u(x_3, t; \bar{K}_{n_1 n_2})_{|x_3 = \alpha_{n_1 n_2} \tilde{S}_{3,0}, t=T}$$

where

$$\bar{K}_{n_1 n_2} := K - e^{L_{11}x_{1,n_1}} - e^{L_{21}x_{1,n_1}+L_{22}x_{2,n_2}},$$
$$\alpha_{n_1 n_2} := e^{L_{31}x_{1,n_1}+L_{32}x_{2,n_2}}$$

and u denotes the solution of (B.4.1).

v) MC1-PDE2 method: Note,

$$P_0 = e^{-rT}E^*\left[\left(K - S_{1,T} - S_{2,T} - S_{1,T}^{-2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}} \tilde{S}_{3,T}\right)^+\right].$$

Simulate N_{MC} realizations of $\tilde{S}_3$. The option price P_0 is then approximated by

$$P_0^a := \frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} u(x_1, x_2, t; \tilde{S}_{3,T}^m)_{|x_1=S_{1,0},x_2=S_{2,0},t=T}$$

where u denotes the solution of the initial value problem for the two dimensional Black-Scholes PDE with parameterized (β) initial condition

$$u_0(x_1, x_2, 0; \beta) = \left(K - x_1 - x_2 - x_1^{-2(L^{-1})_{31}} x_2^{-(L^{-1})_{32}} \beta\right)^+, \quad x_1, x_2 > 0.$$

The problem is

$$\frac{\partial u}{\partial t} - \sum_{i,j=1}^{2} x_i x_j \varrho_{ij} \frac{\partial u}{\partial x_i}\frac{\partial u}{\partial x_j} - r\sum_{i=1}^{2} x_i \frac{\partial u}{\partial x_i} + ru = 0 \qquad \text{in } \Omega \times (0, T) \qquad \text{(B.4.2a)}$$

$$u(t = 0) = u_0 \qquad \text{in } \Omega \qquad \text{(B.4.2b)}$$

where $\Omega = \mathbb{R}_+ \times \mathbb{R}_+$ and

$$\varrho = (\varrho_{ij})_{i,j=1\dots3} = \frac{1}{2}\begin{pmatrix} a_{11}^2 & a_{11}a_{21} \\ a_{11}a_{21} & a_{21}^2 + a_{22}^2 \end{pmatrix}. \qquad \text{(B.4.3)}$$

vi) QUAD1-PDE2 method: Note,

$$P_0 = e^{-rT} \int_{\mathbb{R}} E^*\left[\left(K - S_{1,T} - S_{2,T} - S_{1,T}^{-2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}} e^{x_3}\right)^+\right] f_{3,T}(x_3)dx_3.$$

With the notation from above another approximation of the option price P_0 is

$$P_0^a := \sum_{n=1}^{N_Q} \delta x_3 f_{3,T}(x_{3,n}) e^{-rT} E^*\left[\left(K - S_{1,T} - S_{2,T} - S_{1,T}^{2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}} e^{x_{3,n}}\right)^+\right]$$
$$= \sum_{n=1}^{N_Q} \delta x_3 f_{3,T}(x_{3,n}) u(x_1, x_2, t; x_{3,n})_{|x_1=S_{1,0},x_2=S_{2,0},t=T}$$

where u is the solution of the initial value problem (B.4.2).

vii) MC1-QUAD2 method: Reformulating equation (B.2.2) we deduce

$$P_0 = e^{-rT} E^* \int_{\mathbb{R}^2} \left(K - e^{(Lx)_1} - e^{(Lx)_2} - e^{L_{31}x_1 + L_{32}x_2 + \tilde{s}_{3,T}}\right)^+ f_{1,T}(x_1) f_{2,T}(x_2) dx_1 dx_2$$

and obtain the following method:

Compute N_{MC} realizations of $\tilde{s}_{3,T}$ and approximate P_0 by

$$P_0^a := e^{-rT} \frac{1}{N_{MC}} \sum_{n_1,n_2=1}^{N_Q} \sum_{m=1}^{N_{MC}} \left(\prod_{i=1}^{2} \chi_{n_i} \delta x_i f_{i,T}(x_{i,n_i})\right) \left(K - e^{x_{1,n_1}} - e^{L_{21}x_{1,n_1} + x_{2,n_2}} - e^{L_{31}x_{1,n_1} + L_{32}x_{2,n_2} + \tilde{s}_{3,T}^m}\right)^+ .$$

viii) MC2-QUAD1 method: Note,

$$P_0 = e^{-rT} \int_{\mathbb{R}} E^* \left[\left(K - S_{1,T} - S_{2,T} - S_{1,T}^{-2(L^{-1})_{31}} S_{2,T}^{-(L^{-1})_{32}} e^{x_3}\right)^+\right] f_{3,T}(x_3) dx_3.$$

The method is: simulate N_{MC} realizations of (S_1, S_2) and compute

$$P_0^a := e^{-rT} \frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} \sum_{n=1}^{N_Q} \chi_n \delta x_3 f_{3,T}(x_{3,n}) \left(K - S_{1,T}^m - S_{2,T}^m - {S_{1,T}^m}^{-2(L^{-1})_{31}} {S_{2,T}^m}^{-(L^{-1})_{32}} e^{x_{3,n}}\right)^+ .$$

Combination of three methods

ix) MC1-QUAD1-PDE1 method: Note,

$$P_0 = \int_{\mathbb{R}} f_{2,T}(x_2) e^{-rT} E^* \left[\left(K - e^{\tilde{s}_{1,T}} - e^{L_{21}\tilde{s}_{1,T} + x_2} - e^{(-2(L^{-1})_{31} - (L^{-1})_{32} L_{21})\tilde{s}_{1,T} - (L^{-1})_{32} x_2} \tilde{S}_{3,T}\right)^+\right] dx_2.$$

An approximation to P_0 is then

$$P_0^a := \frac{1}{N_{MC}} \sum_{m=1}^{N_{MC}} \sum_{n=1}^{N_Q} \chi_2 \delta x_2 f_{2,T}(x_{2,n}) u(x_3, t; \bar{K}_n^m)_{|x_3 = \alpha_n^m \tilde{S}_{3,0}, t=T}$$

where

$$\bar{K}_n^m := K - e^{\tilde{s}_{1,T}^m} - e^{L_{21}\tilde{s}_{1,T}^m + x_{2,n_2}}$$
$$\alpha_n^m := e^{(-2(L^{-1})_{31} - (L^{-1})_{32}L_{21})\tilde{s}_{1,T}^m - (L^{-1})_{32}x_{2,n}}$$

and u denotes the solution of (B.4.1).

B.5. Numerical Results

This section provides a documentation of numerical results. We have considered European put options on baskets of three and five assets and used mixed methods to compute their price. If the method is stochastic, i.e. if part of it is Monte Carlo simulation, then we have run the method with different seed values several times (N_S) and computed mean (m) and standard deviation (s) of the price estimates. If the method is deterministic, we have chosen the discretization parameters such that the first three digits of P_0^a remained fix while the discretization parameters have been further refined. Instead solving the one-dimensional Black-Scholes PDE we have used the Black-Scholes formula.

i) European put on three assets: The problem is to compute the price of a European put option on a basket of three asset in the framework outlined in §B.2. We have chosen the parameters as follows: $K = 150$, $T = 1$, $r = 0.05$,

$$\rho = \begin{pmatrix} 1 & -0.1 & -0.2 \\ -0.1 & 1 & -0.3 \\ -0.2 & -0.3 & 1 \end{pmatrix},$$
$$\sigma = \begin{pmatrix} 0.3 & 0.2 & 0.25 \end{pmatrix}^T,$$
$$S_0 = \begin{pmatrix} 55 & 50 & 45 \end{pmatrix}^T.$$

We have used various (mixed) methods to compute approximations to P_0 (see (B.2.2)).
We have used `FreeFem++`[1] and `C++` for the implementation of the methods. If the two-dimensional Black-Scholes PDE had to be solved, we have used `FreeFem++` and otherwise `C++`. The implementation in `FreeFem++` requires a localization and the weak formulation of the Black-Scholes PDE. The triangulation of the computational domain and the discretization of the Black-Scholes PDE by conforming P1 finite elements is done by `FreeFem++`.
A reference result for P_0 has been computed using the Monte Carlo method with 10^7 drawings.

[1] `FreeFem++` is a high level integrated development environment for the numerical solution of partial differential equations. It is a free software based on the Finite Element Method. It has been developed by Frédéric Hecht, Olivier Pironneau and others, see http://www.freefem.org/ff++/ for further information.

The numerical results are displayed in Table B.5.1. One can see that the computational load for the PDE2 methods (MC1-PDE2, QUAD1-PDE2) is much larger than for the other methods. Furthermore the results seem to be less precise than in the other cases. The results have been obtained very fast if just quadrature (QUAD3) or quadrature in combination with the Black-Scholes formula (QUAD2-PDE1) has been used. In these cases the results seem to be very precise although the discretization has been coarse ($N_Q = 12$). Comparison of the results obtained by the MC3 method with the results obtained by the MC2-PDE1 method shows that the last mentioned seems to be superior. The computing time is about equal but the standard deviation is for MC2-PDE1 much less than for MC3.

Method			Parameter				Results		
MC	PDE	QUAD	N_{MC}	N_Q	N_{FE}	N_S	m	s	τ
3	-	-	10^7	-	-	10	3.988	0.002	22.46
3	-	-	25000	-	-	100	3.994	0.046	0.147
2	1	-	25000	-	-	100	3.989	0.029	0.162
1	2	-	100	-	2601	10	3.886	0.195	372.5
-	-	3	-	12	-	-	3.984	-	0.005
-	1	2	-	12	-	-	3.987	-	0.005
-	2	1	-	12	2601	-	4.016	-	42.24
1	-	2	25000	12	-	100	3.991	0.022	2.723
2	-	1	25000	12	-	100	3.987	0.032	0.369
1	1	1	25000	12	-	100	3.990	0.023	0.514

Table B.5.1.: Columns 1-3: the method used to approximate P_0; column 4-7: the discretization parameters: N_{MC} the number of Monte Carlo drawings, N_Q the number subintervals (trapezoidal rule), N_{FE} the number of vertices of the triangulation (finite element method), N_S the number of samples used to compute the mean (m) and the standard deviation (s); columns 8-10: the numerical results: in column 8 the (mean) of P_0^a, in column 9 the standard deviation of P_0^a, in column 10 the computing time (for one sample of P_0^a)

ii) European put on five assets: Let P be a European put option on a basket of five assets. We keep the same notation as introduced in §B.2, i.e. the maturity of the option is T, the strike is K and the payoff function is here

$$\varphi(x) := \left(K - \sum_{i=1}^{5} x_i\right)^+ \qquad x = (x_i)_{i=1}^5 \in \mathbb{R}_+^5.$$

The system of stochastic differential equations which describes the dynamics of the underlying assets has the same form as in (B.2.1). We have set $K = 250$, $T = 1$,

$r = 0.05$,

$$\rho = \begin{pmatrix} 1 & -0.37 & -0.40 & -0.44 & -0.50 \\ -0.37 & 1 & -0.50 & -0.46 & -0.05 \\ -0.40 & -0.50 & 1 & 0.51 & 0.29 \\ -0.44 & -0.46 & 0.51 & 1 & 0.20 \\ -0.50 & -0.05 & 0.29 & 0.20 & 1 \end{pmatrix},$$
$$\sigma = \begin{pmatrix} 0.3 & 0.275 & 0.25 & 0.225 & 0.2 \end{pmatrix}^T,$$
$$S_0 = \begin{pmatrix} 40 & 45 & 50 & 55 & 60 \end{pmatrix}^T.$$

We approximated the price of P at time 0 by various (mixed) methods. The results are displayed in Table B.5.2. One can see that for all tested methods the (mean) price has been close (± 0.003) to the reference price (1.159). Since $N_Q = 10$ turned out to be enough the computational effort has been very low for QUAD5 and QUAD4-PDE1. In the case the method is stochastic, deterministic methods allowed to reduce the variance, cf. MC4-QUAD1 and MC4-PDE1-QUAD1.

Method			Parameter			Results		
MC	PDE	QUAD	N_{MC}	N_Q	N_S	m	s	τ
5	-	-	10^7	-	10	1.159	0.001	27.67
5	-	-	25000	-	100	1.161	0.019	0.162
4	-	1	25000	-	100	1.156	0.015	0.174
-	-	5	-	10	-	1.161	-	0.082
-	1	4	-	10	-	1.159	-	0.036
3	1	1	25000	10	100	1.158	0.013	0.442

Table B.5.2.: Columns 1-3: the method used to approximate P_0; column 4-6: the discretization parameters: N_{MC} the number of Monte Carlo drawings, N_Q the number subintervals (trapezoidal rule), N_S the number of samples used to compute the mean (m) and the standard deviation (s); columns 7-9: the numerical results: in column 7 the (mean) of P_0^a, in column 8 the standard deviation of P_0^a, in column 9 the computing time (for one sample of P_0^a)

B.6. Conclusion

We have proposed to combine stochastic (Monte Carlo) and deterministic (quadrature and PDE) methods for the valuation of European options on a basket of assets.

We have chosen a basket of three assets and a standard payoff function for the presentation of the methods. Numerical results have been provided for a basket of three and five assets.

Other payoff functions may require some modifications in the discretization process. We have not investigated the case when volatilities are time and/or space depended. This could be future work.

The numerical results have shown that the combinations of the Monte Carlo method with quadrature and/or the Black-Scholes formula is advantageous. Using quadrature and the Black-Scholes formula reduces the computational load. Compared to standard Monte Carlo, the speed-up can be of factor 50 (see Table B.5.1). Further numerical experiments are, however, required to see whether quadrature is still competitive if the payoff function is more complicated. In this case precise results may only be obtained if N_Q is chosen larger.

List of Figures

List of Tables

Bibliography

[AAE02] L. Andersen, J. Andreasen, and D. Eliezer. Static replication of barrier options: some general results. *J. Comput. Finance*, 5(4):1–25, 2002. 4.1

[AHHS10] Harbir Antil, Matthias Heinkenschloss, Ronald Hoppe, and Danny Sorensen. Domain decomposition and model reduction for the numerical solution of pde constrained optimization problems with localized optimization variables. *Computing and Visualization in Science*, pages 249–264, June 2010. (document)

[Ant05] A. Antoulas. *Approximation of Large-Scale Dynamical Systems*. Advances in Design and Control. Society for Industrial Mathematics, 1 edition, June 2005. (document)

[AP05] Y. Achdou and O. Pironneau. *Computational Methods for Option Pricing*. Frontiers in Applied Mathematics. SIAM, 2005. 2, 4.1, 4.2, 4.3

[AS64] Milton Abramowitz and Irene A. Stegun. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables*. Dover, New York, ninth dover printing, tenth gpo printing edition, 1964. A.1.1, A.1.1

[BB02] M. Breitner and T. Burmester. Optimization of european double-barrier-options via optimal control of the black-scholes-equation. In P. Chamoni et al., editor, *Operations Research Proceedings '01*. Springer, Berlin-Heidelberg-New York, 2002. 4.1, 4.2

[BBG97] Phelim Boyle, Mark Broadie, and Paul Glasserman. Monte carlo methods for security pricing. *Journal of Economic Dynamics and Control*, 21(8-9):1267–1321, 1997. 3.6, (document)

[BC94] Jonathon Bowie and Peter Carr. Static simplicity. *Risk 7*, pages 45–49, 1994. 2.1

[BE80] Phelim P. Boyle and David Emanuel. Discretly adjusted option hedges. *Journal of Financial Economics*, pages 259–282, 1980. 1.1

[BK04] Nicholas H. Bingham and Rüdiger Kiesel. *Risk-Neutral Valuation: Pricing and Hedging of Financial Derivatives*. Springer, 2004. 3.3.1

[Boy77] Phelim P. Boyle. Options: A monte carlo approach. *Journal of Financial Economics*, 4(3):323–338, May 1977. B.1

[BP03] Jean-Philippe Bouchaud and Marc Potters. *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management.* Cambridge University Press, 2003. 2.2.1

[BS72] Fischer Black and Myron Scholes. The valuation of option contracts and a test of market efficiency. *The Journal of Finance*, 27(2):399–417, May 1972. 1.1

[BS73] Fischer Black and Myron Scholes. The pricing of options and corporate liabilities. *Journal of Political Economy*, 81:637–654, 1973. 1

[BS94] Jean-Philippe Bouchaud and Didier Sornette. The black-scholes option pricing problem in mathematical finance: Generalizations and extensions for a large class of stochastic processes. *Journal de Physique I*, 4:863–881, 1994. 2.2.1

[BSG05] Sebastien Bossu, Eva Strasser, and Regis Guichard. Just what you need to know about variance swaps. Technical report, JPMorgan, February 2005. 3.1

[BW07] Lewis Biscamp and Tim Weithers. Variance swaps and cboe s&p 500 variance futures. Technical report, Chicago Trading Company, LLC, 2007. Euromoney handbooks. 3.1

[Car95] P. Carr. Two extensions to barrier option pricing. *Appl. Math. Finance*, 2:173–209, 1995. 4.1

[CC97] P. Carr and A. Chou. Breaking barriers. *Risk Magazine*, 10:139–144, 1997. 4.1

[CEG98] P. Carr, K. Ellis, and V. Gupta. Static hedging of exotic options. *J. Finance*, 53:1165–1190, 1998. 4.1

[CL09] Peter Carr and Roger Lee. Volatility derivatives. *Annual Review of Financial Economics*, 1:1–21, 2009. 3.1

[CLL07] Thomas F. Coleman, Dmitriy Levchenkov, and Yuying Li. Discrete hedging of american-type options using local risk minimization. *Journal of Banking & Finance*, 31(11):3398–3419, November 2007. (document)

[CLP10] Rama Cont, Nicolas Lantos, and Olivier Pironneau. A reduced basis for option pricing. HAL, September 2010. hal.archives-ouvertes.fr. (document)

[CS97] Les Clewlow and Chris Strickland. *Exotic options the state of the art.* International Thomson Business Press, March 1997. 2

[CV96] T.H.E. Cheuk and T.C.F. Vorst. Complex barrier options. *The Journal of Derivatives*, 4:8–22, 1996. 4.1

[DDKZ99] Kresimir Demeterfi, Emanuel Derman, Michael Kamal, and Joseph Zou. More than you ever wanted to know about volatility swaps. Quantitative strategy research notes, Goldman Sachs, March 1999. 3.1

[DEK95] Emanuel Derman, Deniz Ergener, and Iraj Kani. Static options replication. *The Journal of Derivatives*, 2(4):78–95, 1995. 2.1, 4.1

[Dev86] Luc Devroye. *Non-Uniform Random Variate Generation*. Springer Verlag Gmbh, 1986. 1.4.1

[DH08] Peter Deuflhard and Andreas Hohmann. *Numerische Mathematik 1: Eine algorithmisch orientierte Einführung*. Gruyter, 2008. A.1.2, A.1.2

[Fin03] Jason Fink. An examination of the effectiveness of static hedging in the presence of stochastic volatility. *The Journal of Futures Markets*, 23(9):859–890, 2003. 2.1

[Fis96] George S. Fishman. *Monte Carlo: Concepts, algorithms, and applications*. Springer Series in Operations Research. Springer-Verlag, New York, 1996. 3.5.1

[FS86] Hans Föllmer and Dieter Sondermann. Hedging of non-redundant contingent claims. In Werner Hildenbrand and Andreu Mas-Colell, editors, *Contributions to Mathematical Economics: In Honor of Gérard Debreu*, chapter 12, pages 205–223. North-Holland, 1986. 1, 1.1

[FS04] Hans Föllmer and Alexander Schied. *Stochastic Finance. An Introduction in Discrete Time*. Gruyter, 2004. 1.2, 1.2, 1.2

[Gla03] Paul Glasserman. *Monte Carlo Methods in Financial Engineering*, volume 53 of *Applications of Mathematics*. Springer, August 2003. 3.5.1, 3.6, (document)

[GM07] A.M. Giese and J.H. Maruhn. Cost-optimal static super-replication of barrier options - an optimization approach. *J. Comput. Finance*, 10:71–97, 2007. 4.1

[GMW82] Philip E. Gill, Walter Murray, and Margaret H. Wright. *Practical optimization*. Emerald Group Pub, revised edition, Januar 1982. 4.1

[Gri00] Andreas Griewank. *Evaluating derivatives: principles and techniques of algorithmic differentiation*, volume 19 of *Frontiers in applied mathematics*. SIAM, 2000. 4.1

[Hes93] Steven L. Heston. A closed-form solution for options with stochastic volatility with applications to bond and currency options. *The Review of Financial Studies*, 6(2):327–343, 1993. 1.1

[HK79] J. Michael Harrison and David M. Kreps. Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory*, 20(3):381–408, June 1979. 1

[HK94] R. Heynen and H. Kat. Crossing barriers. *Risk Magazine*, 7:46–51, 1994. 4.1

[HL11] Ronald H.W. Hoppe and Tobias Lipp. Optimal control of European double barrier basket options. *Journal of Numerical Mathematics*, 19(2):113–136, 2011. 4

[HP81] J. M. Harrison and S. R. Pliska. Martingales and stochastic integrals in the theory of continuous trading. *Stochastic Processes and their Application*, 3(11):215–260, 1981. 1

[HPS01a] David Heath, Eckhard Platen, and Martin Schweizer. A comparison of two quadratic approaches to hedging in incomplete markets. *Mathematical Finance*, 11(4):385–413, 2001. 1.1

[HPS01b] David Heath, Eckhard Platen, and Martin Schweizer. Numerical comparison of local risk-minimisation and mean-variance hedging. In Elyès Jouini, Jaksa Cvitanic, and Marek Musiela, editors, *Handbooks in Mathematical Finance: Option Pricing, Interest Rates and Risk Management*, chapter 14, pages 509–537. Cambridge University Press, 2001. 1.1

[HPUU08] Michael Hinze, Rene Pinnau, Michael Ulbrich, and Stefan Ulbrich. *Optimization with PDE Constraints*, volume 23 of *Mathematical Modelling: Theory and Applications*. Springer, 1 edition, November 2008. 2

[HRA11] M. Heinkenschloss, T. Reis, and A.C. Antoulas. Balanced truncation model reduction for systems with inhomogeneous initial conditions. *Automatica*, 47:559–564, 2011. (document)

[Hul08] John Hull. *Options, Futures, and Other Derivatives.* Prentice Hall International, 2008. 4.1, A.3

[Jae02] Peter Jaeckel. *Monte Carlo Methods in Finance.* Wiley, April 2002. 3.5.1

[JYC03] M. Jeanblanc, M. Yor, and M. Chesney. *Mathematical Models for Financial Markets.* Springer, Berlin-Heidelberg-New York, 2003. 2

[Kel99] C.T. Kelley. *Iterative Methods for Optimization.* SIAM, Philadelphia, 1999. 4.7

[KP92] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations.* Application of Mathematics. Springer, 1992. 2.2.1

[KW08] Malvin H. Kalos and Paula A. Whitlock. *Monte Carlo Methods*. Wiley-VCH Verlag GmbH & Co. KGaA, 2 edition, 2008. 1.4.1

[KWY98] Y.-K. Kwok, L. Wu, and H. Yu. Pricing multi-asset options with an external barrier. *Int. J. Theoretical Appl. Finance*, 1:523–541, 1998. 4.1

[Lio71] Jacques Louis Lions. *Optimal Control Of Systems Governed By Partial Differential Equations*. Springer-Verlag, New York, 1971. 2

[LMT01] Jacques-Louis Lions, Yvon Maday, and Gabriel Turinici. Résolution d'edp par un schéma en temps pararéel - a parareal in time discretization of pde's. *Comptes Rendus De L Academie Des Sciences Serie I-mathematique*, 332:661–668, 2001. (document)

[LS98] Francis A. Longstaff and Eduardo S. Schwartz. Valuing american options by simulations: A simple least-squares approach. *Working Paper Anderson Graduate School of Management University of California*, 25, 1998. (document), 1, 1.1, 2.3.2

[LYH03] C.F. Lo, P.H. Yuen, and C.H. Hui. A simple approach for pricing barrier options with time-dependent parameters. *Quantitative Finance*, 3:98–107, 2003. 4.1

[Mar09a] Jan H. Maruhn. *Robust static super-replication of barrier options*. Radon Series on Computational and Applied Mathematics. de Gruyter, 2009. 2.1

[Mar09b] J.H. Maruhn. Duality in static hedging of barrier options. *Optimization*, 58:319–333, 2009. 4.1

[McL05] Don McLeish. Probability. http://sas.uwaterloo.ca/ dlmcleis/s901/, 2005. 1.3

[Mer73] Robert C. Merton. Theory of rational option pricing. *Bell Journal of Economics and Management Science*, 4:141–183, 1973. 4.1

[MN01] Manuel Moreno and Javier R. Navas. On the robustness of least-squares monte carlo (lsm) for pricing american derivatives. Economics Working Papers 543, Department of Economics and Business, Universitat Pompeu Fabra, Apr 2001. 2.3.2

[MPT02] Yvon Maday, Anthony T. Patera, and Gabriel Turinici. A priori convergence theory for reduced-basis approximations of single-parametric elliptic partial differential equations. *Journal of Scientific Computing*, 17:437–446, 2002. (document)

[MS09] J.H. Maruhn and E.W. Sachs. Robust static hedging of barrier options in stochastic volatility models. *Mathematical Methods of Operation Research*, 70:405–433, 2009. 4.1

[MT02] Yvon Maday and Gabriel Turinici. A parareal in time procedure for the control of partial differential equations. *Comptes Rendus Mathematique*, 335(4):387–392, 2002. (document)

[Nel99] I. Nelken. *Pricing, Hedging, and Trading Exotic Options*. McGraw-Hill, New York, 1999. 2

[NP06] M. Nalholm and R. Poulsen. Static hedging of barrier options under general asset dynamics: unification and application. *J. Derivatives*, 13:46–60, 2006. 4.1

[NW99] J. Nocedal and S.J. Wright. *Numerical Optimization*. Springer, Berlin-Heidelberg-New York, 2nd edition, 1999. 4.1, 4.7

[PB03] Benoît Pochart and Jean-Philippe Bouchaud. Option pricing and hedging with minimum local expected shortfall. arXiv, August 2003. (document)

[PBS01] Marc Potters, Jean-Philippe Bouchaud, and Dragan Sestovic. Hedged monte-carlo: low variance derivative pricing with objective probabilities. *Physica A*, 289:517–525, 2001. 1.1, 1.1, (document)

[Pir09] Olivier Pironneau. Calibration of options on a reduced basis. *Journal of Computational and Applied Mathematics*, 232(1):139–147, October 2009. (document)

[Pou06] R. Poulsen. Barrier options and their static hedges: simple derivations and extensions. *Quant. Finance*, 6:327–335, 2006. 4.1

[QV99] Alfio Quarteroni and Alberto Valli. *Domain Decomposition Methods for Partial Differential Equations*. Numerical Mathematics and Scientific Computing. Oxford University Press, July 1999. (document)

[Rit98] P. Ritchken. On pricing barrier options. *Currency derivatives: pricing theory, exotic options, and hedging applications*, pages 275–289, 1998. 4.1

[RR91] E. Reiner and M. Rubinstein. Breaking down the barriers. *Risk Magazine*, 4:28–35, 1991. 4.1

[RR93] M. Renardy and R.C. Rogers. *An Introduction to Partial Differential Equations*. Springer, Berlin-Heidelberg-New York, 1993. 4.3

[Sbu05] A. Sbuelz. Hedging double barriers with singles. *Int. J. Theoretical and Applied Finance*, 8:393–406, 2005. 4.1

[Sch84] Martin Schweizer. Varianten der Black-Scholes-Formel. Diploma thesis, Eidgenössische Technische Hochschule Zürich, 1984. 1

[Sch88] Martin Schweizer. *Hedging of Options in a General Semimartingale Model*. Phd thesis, Eidgenössische Technische Hochschule Zürich, 1988. 1, 1.1

[Sch90] Martin Schweizer. Risk-minimality and orthogonality of martingales. *Stochastics and Stochastics Reports*, 30:123–131, 1990. 1

[Sch91] Martin Schweizer. Option hedging for semimartingales. *Stochastic Processes and their Applications*, 37(2):339–363, 1991. 1

[Sch08] Martin Schweizer. Local risk-minimization for multidimensional assets and payment streams. *Banach Center Publications*, 83:213–229, 2008. (document)

[SH88] M. Schweizer and Föllmer H. Hedging by sequential regression: An introduction to the mathematics of option trading. *ASTIN Bulletin*, 18(2):147–160, 1988. 1, 1.1

[Sie06] Kunibert G. Siebert. Numerische Mathematik I Vorlesung im Wintersemester 2006/2007. http://www.mathematik.uni-stuttgart.de/fak8/ians/lehrstuhl/nmh/index.html, 2006. 2.3.2

[SS91] Elias M. Stein and Jeremy C. Stein. Stock price distribution with stochastic volatility: An analytic approach. *Review of Financial Studies*, 4(4):727–752, 1991. 1.1

[SS08] Ekkehard Sachs and M. Schu. Reduced order models (pod) for calibration problems in finance. In Karl Kunisch, Günther Of, and Olaf Steinbach, editors, *Numerical Mathematics and Advanced Applications*, Proceedings of ENUMATH 2007, pages 735–742. Springer, 2008. (document)

[Top98] J. Topper. Finite element modeling of exotic options. Discussion Paper 216, Arthur Andersen Risk Management,, Eschborn, 1998. 4.2

[Trö10] Fredi Tröltzsch. *Optimal Control of Partial Differential Equations. Theory, Methods, and Applications.*, volume 112 of *Graduate Studies in Mathematics*. American Mathematical Society, Providence, 2010. 2

[TW04] Andrea Toselli and Olof Widlund. *Domain Decomposition Methods - Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer, Berlin, September 2004. (document)

[Vol] Stefan Volkwein. Model reduction using proper orthogonal decomposition. www.uni-graz.at/imawww/volkwein/POD.pdf. (document)

[WHD95] Paul Wilmott, Sam Howison, and Jeff Dewynne. *The Mathematics of Financial Derivatives: A Student Introduction*. Cambridge University Press, 1995. A.3

[Wil07] Paul Wilmott. *Paul Wilmott Introduces Quantitative Finance*. John Wiley & Sons, June 2007. (document), 2.1, A.3

[WK03] H.Y. Wong and Y.-K. Kwok. Multi-asset barrier options and occupation time derivatives. *Appl. Math. Finance*, 10:245–266, 2003. 4.1

[Zha97] P.G. Zhang. *Exotic Options*. World Scientific, Singapore, 1997. 2

Numerical methods for optimization in finance: optimized hedges for options and optimized options for hedging

Abstract: This dissertation contributes to optimization in finance through numerical methods. The input consists of two parts:

In part 1, we propose a numerical method to compute a trading strategy for the hedging of a financial derivative with N hedging instruments. The underlying mathematical framework is local risk minimization in discrete time. The method combines Monte Carlo simulation with least squares regression in analogy to the method of Longstaff and Schwartz. We study the proposed method on two example problems. For both problems the number of hedging instruments is two. One of the hedging instruments is always the underlying asset of the hedging objective. The other hedging instrument is a vanilla put option in the first example and a variance swap in the second example.

In part 2, we propose an optimal control approach for the optimization of European double barrier basket options. The basket consists of two assets. The objective is to control the payoff and the rebate at the upper barrier such that the delta of the option is as close as possible to a predefined constant. This gives rise to a control constrained optimal control problem for the (two-dimensional) Black-Scholes equation with Dirichlet boundary control and finite time control. Based on the variational formulation of the problem in an appropriate Sobolev space setting, we prove the existence of a unique solution and state the first order necessary optimality conditions. Discretization in space by P1 finite elements and discretization in time by the backward Euler scheme results in a fully discrete optimal control problem. Numerical results illustrate the benefits optimized double barrier options.

Numerische Verfahren zur Optimierung in der Finanzwirtschaft: optimierte Absicherungsstrategien für Optionen und optimierte Optionen zur Absicherung

Zusammenfassung: Diese Dissertation trägt mittels numerischer Verfahren zur Optimierung in der Finanzwirtschaft bei. Die Dissertation besteht aus zwei Teilen.

Im ersten Teil wird ein numerisches Verfahren zur Berechnung von Handelsstrategien vorgestellt. Mit den Strategien lassen sich Derivate mit N Absicherungsinstrumenten absichern. Die zugrunde liegende mathematische Theorie ist die der lokalen Risikominimierung in diskreter Zeit. Das Verfahren basiert auf Monte-Carlo-Simulation und der Methode der kleinsten Quadrate und hat Ähnlichkeit zum Verfahren von Longstaff und Schwartz. Wir studieren das Verfahren an zwei Beispielen. In beiden Beispielen gibt es zwei Absicherungsinstrumente wobei eines davon das dem Derivat zugrunde liegende Gut ist. Das andere Absicherungsinstrument ist im ersten Beispiel eine Standardverkaufsoption und im zweiten Beispiel ein Varianzswap.

Im zweiten Teil wird ein Optimalsteuerungsansatz zur Optimierung von Europäischen Korboptionen mit doppelter Schwelle vorgestellt. Der Korb besteht aus zwei Werten. Das Ziel ist die Auszahlung an der oberen Schwelle und die Auszahlung bei Fälligkeit derart zu steuern, dass das Delta der Option so nahe wie möglich zu einer a priori festgelegten Konstanten ist. Das führt auf ein steuerungsbeschränktes Optimalsteuerungsproblem für die zwei-dimensionale Black-Scholes Gleichung mit Dirichlet Randsteuerung und Endzeitsteuerung. Basierend auf der Variationsformulierung des Problems in passenden Sobolev-Räumen zeigen wir die Existenz und Eindeutigkeit der Lösung. Zudem leiten wir notwenidge Optimalitätsbedingungen erster Ordnung her. Diskretisierung im Raum mit P1 Finiten Elementen und Diskretisierung der Zeit mit dem impliziten Euler-Schema führt auf ein volldiskretes Optimalsteuerungsproblem. Numerische Ergebisse zeugen von den Vorzügen optimierter Europäischer Korboptionen mit doppelter Schwelle.

Méthodes numériques pour l'optimisation en finance : couverture optimisée pour des options et options optimisées pour la couverture

Résumé : Cette thèse porte sur l'optimisation en finance par des méthodes numériques. La thèse se présente en deux parties.

Dans la première partie, nous proposons une méthode numérique pour calculer une stratégie de trading pour la couverture d'un produit financier dérivé avec plusieurs instruments de couverture. Le cadre mathématique sous-jacent est la minimisation du risque local en temps discret. La méthode combine la simulation de Monte-Carlo et la régression des moindres carrés - analogue à la méthode de Longstaff et Schwartz. Nous l'appliquons à deux exemples particuliers. Les instruments de couverture sont l'actif sous-jacent, des options vanilles et des swaps de variance.

Dans la seconde partie, nous proposons une approche par contrôle optimal pour l'optimisation des options paniers à barrière double de type européen. Le panier est constitué de deux actifs. L'objectif est de contrôler le versement à la barrière supérieure et le versement à la date d'échéance de sorte que le delta de l'option soit aussi proche que possible d'une constante prédéfinie. Cela donne lieu à un problème de contrôle optimal de type contrôle restreint pour l'équation aux dérivées partielles de Black-Scholes avec des conditions de Dirichlet au bord contrôlées et de condition terminale contrôlée. En utilisant la formulation variationnelle du problème dans un cadre d'espace de Sobolev à poids, on prouve l'existence et l'unicité de la solution. Les discrétisations par la méthode des éléments finis et par le schéma d'Euler implicite conduisent à un problème de contrôle optimal entièrement discret. Des résultats numériques pour des problèmes de test sélectionnés illustrent les avantages de couvrir avec des options paniers à barrière double de type européen dans le cas des versements contrôlés optimals.

Lebenslauf

Name:	Tobias Lipp
Geboren:	16. April 1984 in Lauingen (Donau)
Staatsangehörigkeit:	Bundesrepublik Deutschland

Wissenschaftlicher Werdegang:

10/2003-05/2008	Studium an der Universität Augsburg (Abschluss: Diplom Mathematik)
08/2006-08/2007	Studium an der Universität Houston (Abschluss: Master of Science Mathematics)
10/2008-05/2012	Promotionsstudium an den Universitäten Augsburg und Pierre und Marie Curie - Paris 6